Inference Methods

for

Gravitational Wave Data Analysis

On the application of Bayesian inference and modern modelling methods to astrophysical problems in the era of gravitational wave observation.

Daniel Williams

November 2019

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy.

School of Physics & Astronomy

College of Science and Engineering

University of Glasgow · Oilthigh Glaschu

Abstract

Einstein's publication of the *general theory of relativity* in 1915, and the discovery of a wave-like solution to the field-equations of that theory sparked a century-long quest to detect *gravitational waves*. These illusive metric disturbances were predicted to ripple-away from some of the most energetic events in the universe, such as supernovae and colliding black holes.

The quest was completed in September 2015, with the Laser interferometer gravitational-wave observatory (LIGO) observation of a gravitational wave produced by a pair of coalescing black holes, but work to continue detecting and interpretting the signals which are detected by LIGO and its brethren is by no means complete. The age of gravitational wave observation has arrived, and with it the difficulties of interpretting myriad signals, differentiating them from noise, and analysing them in order to gain insight into the astrophysical systems which produced them.

This thesis provides overview of the history of the field of gravitational wave science: both in terms of the theoretical principles which frame it, and the attempts to build instruments which could measure them. It then provides a discussion of the morphologies of the signals which are searched for in current detectors' data, and the astrophysical systems which may produce such signals. It is of great importance that the sensitivity of both detectors and the signal analysis techniques which are used is well-understood. A substantial part of the novel work presented in this document discusses the development of a technique for assessing this sensitivity, through a software package called Minke.

Knowing the sensitivity of a detector to signals from an astrophysical source allows robust limits to be placed on the rate at which these events occur. These rates can then be used to infer properties

of astrophysical systems; this document contains a discussion of a technique which was developed by the author to allow the determination of the geometry of beamed emission from short gamma ray bursts which result from neutron star coalescences. This method finds that at its design sensitivity we expect the advanced LIGO detector to be able to place limits on the opening angle, θ, of the beam within $\theta \in (8.10^{\circ}, 14.95^{\circ})$ under the assumption that all neutron star coalesences produce jets, and that gamma ray bursts occur at an illustrative rate of $\mathcal{R}_{\mathrm{grb}} = 10\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$.

The most efficient methods for extracting signals from noisy data, such as that produced by gravitational wave detectors, and then analysing these signals, requires robust prior knowledge of the signals' morphologies. The development of a new model for producing gravitational waveforms for coalescing binary black hole systems is discussed in detail in this work. The method which is used, Gaussian process regression, is introduced, with an overview of different methods for implementing models which use the method. The model, named Heron, is itself presented, and comparisons between the waveforms produced by Heron and other models which are currently used in analysis are made. Comparisons between the Heron model and highly accurate numerical relativity waveforms are also shown.

This document has been assigned the identifier LIGO-P1900001 by the LIGO Document Control Center.

Contents

List of Figures

List of Tables

Acknowledgements

> "Any road followed precisely to its end leads precisely nowhere.
> Climb the mountain just a little bit to test that it's a mountain.
> From the top of the mountain, you cannot see the mountain."
> — Frank Herbert, *Dune*.

The last three-and-a-bit (well, let's be honest, almost four) years have been pretty busy around the Institute for Gravitational Research (IGR). There have been a few events of some note over that time, and it's been very exciting to be around for them, starting out with what felt like a court-side seat for the adrenaline-fuelled months following the first detection in September 2015, through to being part of the team now that we've entered the era of "routine detection".

So my first thanks go to all of the members of the IGR, who kindly allowed me to spend that time working with them, and learning from them. More generally, I extend my thanks to the members of the LIGO Scientific Collaboration who mentored me as I started to find my way into the field.

I had plenty of opportunity to see parts of the world I'd never have expected to see before starting on this endeavour, and my thanks go to all of the kind hosts who looked after me, whether it was at a world-class lab in the middle of a swamp in Louisiana; in humid Midtown, Atlanta; or the entirely exciting Daejong.

Without the people around me to keep me sane, or distract me from my work, I doubt this tome would ever have come about. Euan, Magnus, Rhys, Shona, Laura, Alex, and David; you all bear some responsibility for this happening (and must accept some thanks and credit for the fact that it did). There were times when I doubted I'd finish at all, and without Andrew

I'm sure I wouldn't: thanks for being around to listen to me complaining, and for letting yourself get embroiled in some of my harebrained schemes.

Then to my supervisors, without whom I'd have never discovered the joys of the Bayesian (correct) interpretation of probability: Siong and Graham, thank you for taking on the risk of taking me on as a student. My examiners, Iain Hannah and Stephen Fairhurst, also deserve due credit for spotting a veritable ocean of typos, and a suggesting a number of edits which greatly aided in improving the clarity of this thesis.

Finally, without my family this could not have happened. Thank you Zoe, not least for all of the funny animal videos and pictures. And thank you to my parents, who first showed me the stars.

My time as a PhD student, and my research, was kindly supported by the Science and Technology Facilities Council (STFC).

Declaration

The material presented in this thesis is the result of my own work in the Institute for Gravitational Research in the School of Physics and Astronomy at the University of Glasgow. It has not been submitted for any other degree at the University of Glasgow, or at any other institution.

This work was carried-out within the LIGO Scientific Collaboration, and as such was facilitated through interaction and collaboration with many other researchers. The opening paragraphs of each chapter in this thesis contain a detailed description of the origin of the work in each chapter. Chapters 5 and 7 summarise two publications: the former has previously been published in *The Astrophysical Journal* [1], and the latter has been submitted for publication, and is available on the *ArXiV* [2].

Acronyms

ASD amplitude spectral density 14, 15, 29, 30, 33, 76, 223–226

BAM bi-functional adaptive mesh 62

BATSE Burst and Transient Source Experiment 123, 125

BBH binary black hole 9, 26, 42–44, 55, 58–65, 67, 68, 77, 82, 89, 109, 110, 113, 187, 197, 205, 219, 221, 224, 225, 230

BNS binary neutron star viii, xi, 4, 44, 46, 55, 60, 109, 110, 112, 120, 124–132, 134–138, 140, 143–151, 220, 231

BSSN Baumgarte-Shapiro-Shibata-Nakamura 62

BW Bayeswave 87, 93–95

CBC compact binary coalescence 28, 38, 42, 45, 51, 69, 72, 93, 130, 188

CCSN core-collapse supernova 73, 74, 123, 124

CGRO Compton Gamma-Ray Observatory 123

CMB cosmic microwave background 54

cWB Coherent WaveBurst 42, 87, 93, 95

DAG directed acyclic graph 112

DECIGO Deci-hertz Interferometer Gravitational wave Observatory 16, 28–30, 41, 73, 74, 76, 78

EFE Einstein field equations 61

EM electromagnetic 126, 151, 220, 221

EOB effective one body 67, 68

EPTA European Pulsar Timing Array 17

ET Einstein Telescope 76, 78

LSC LIGO Scientific Collaborations 232

LVC LIGO and Virgo Scientific Collaborations 3

MAP maximum a posteriori estimate 138, 144, 181

MCMC Markov chain Monte Carlo 119, 144, 171

MDC mock data challenge 90–92, 94, 97, 223

MWEG Milky Way equivalent galaxy 132

NR numerical relativity 60–65, 67–69, 76, 91, 187–190, 196–210, 214, 219, 220, 234

NS neutron star 52, 53

NSBH neutron star / black hole 89, 124, 125

oLIB Omicron LIB 87, 93, 95

PDF probability density function 104, 105

PN post-Newtonian 59, 63–65, 67, 76

PNS proto-neutron star 73, 74

PPTA Parkes Pulsar Timing Array 17

PSD power spectral density 10, 11, 32, 34, 35, 37, 39, 89

RBF radial basis function 162, 166

RMSE root mean squared error 178, 179

RSM response surface methodology 156

SE squared exponential 163–165, 170, 190, 191, 195

sGRB short gamma ray burst viii, xi, 85, 109, 110, 112, 113, 120, 123–131, 133, 134, 137–151, 220, 221

SMBBH supermassive binary black hole 58

SN supernova 71, 93, 124

SNR signal-to-noise ratio 10, 11, 28, 43, 76, 77, 79, 130, 131, 193, 202, 203, 207, 225, 226, 231

SpEC Spectral Einstein Code 62, 63

SVI stochastic variational inference 176, 177

SVR support vector machine regression 156

Notational conventions

Throughout this work I take the convention that the metric tensor, g should be positive, having the signature $(-,+,+,+)$, and likewise the Riemann, R, and Einstein, E, tensors should also be positive, following the "spacelike convention" of Landau & Lifshitz [3], and the convention of Misner, Thorne, and Wheeler [4]. I also adopt the convention of using Greek indices for four-dimensional tensor quantities, such as 4-vectors, and latin indices otherwise. The reader should note that while the discussion of metrics in the context of general relativity is limited to four-dimensions, those metrics used in feature-space descriptions of data, especially in the context of Gaussian process regression, are not. In chapter 1, and in places where general relativity is discussed it should be assumed that indexed tensor quantities follow the Einstein summation convention over repeated indices.

1 Gravitational Waves: generation, propagation, and detection

This chapter introduces gravitational waves, and briefly discusses their description in general relativity, and their propagation through spacetime. The chapter then continues to discuss methods by which they might be detected, and a brief history of attempts to do so.

Section 1.1 contains an introduction and overview of the radiation predicted by general relativity: gravitational waves, and gives a very concise introduction to how these waves propagate. Section 1.2 briefly summarises a number of conventions which are used within gravitational wave astronomy, and this work, for describing the strain effect of gravitational waves, and related quantities. Section 1.3 gives an overview of the history of gravitational wave detection, including the early attempts to identify measurable quantities from the theory, to the operational constraints of modern detectors. Section 1.4 discusses some of the major sources of noise which are introduced into detector data by environmental and instrumental phenomena. Section 1.5 provides an overview of the present network of gravitational wave detectors around the world. Section 1.6 is a short discussion of the detections which have been made to date in the

first two observing runs of the advanced-era detectors.

The vast majority of material in this chapter is review material, however a number of sensitivity curve plots were produced by a Python package which I developed, `gravpy`, which is described in appendix A. This is noted in the captions for these figures.

---

14 September 2015 will likely be remembered as one of the most significant in the history of astronomy, and of astrophysics. Early in the morning of this autumn day, at 09:50 UTC, a gravitational wave (GW) passed through the Earth, and on its way produced a sufficiently large movement in the mirrors and test masses of the detectors of Laser interferometer gravitational-wave observatory (LIGO), as to be detected.

Over five months of data analysis, detector characterisation, and detection verification were conducted by a global team of scientists, in the LIGO and Virgo Scientific Collaborations (LVC). This process resulted in a slew of journal papers being written, vast quantities of data produced, and the launch of an enormous public outreach effort. Eventually, the collaboration found itself in a position to make the announcement of the first direct detection of GWs, GW150914, on 11 February 2016.

Further discoveries were made during the first observing run of the advanced LIGO facilities, LVT151012 [1], observed in October, and GW151226, observed in December, added to the collection of directly detected black hole binaries (although the significance of the former was insufficient to garner its own press release and announcement publication). The majority of 2016 was dedicated to introducing upgrades to the two LIGO interferometers, and to commissioning a third second-generation interferometric observatory, Virgo, located in Italy. The second observation run of Advanced LIGO, which was scheduled to be longer and more sensitive than its first lead to the observation of three further binary black hole coalescences during the run: GW170104, GW170608, and GW170814, the latter observed by a network of three detectors, Virgo having joined

[1]The designation "LVT", or "LIGO / Virgo transient" was used during the first two observing runs for events which were significant, but which did not surpass a threshold of 5σ for that significance. This event was eventually upgraded to the status of a confident event with the publication of the second observing run results [5], and is now known as GW151012.

the observing run a short time before. These events are summarised in table 1.6.1, alongside detections which were made "offline" in the archival data.

Great excitement was also to be had during the second observing run, with the detection of GW170817, the first detection of a coalescing binary neutron star (BNS) system, which was also the first occasion on which gravitational waves had contributed to a multi-messenger observation. This event is discussed in greater detail in section 1.6.2.

GW150914 is not the first evidence for GWs, with the Hulse-Taylor pulsar [6, 7] having provided compelling indirect evidence which lead to its discoverers receiving the 1993 Nobel Prize.

This chapter serves as a brief overview of the mathematical and physical motivation for the search for GWs, the current network of GW detectors, and a synopsis of the first GW detections. In addition a primer on the conventions used in the description of GW signals is included later in this chapter (in section 1.2).

1.1 | Gravitational waves and general relativity

GWs are one of the predictions of Einstein's 1915 General Theory of Relativity [8], with the first theoretical prediction of their existence being proposed by Einstein in 1916 [9, 10], less than a year after the publication of the general theory. Einstein's original wave solution proposes three different forms of GW, which Hermann Weyl denoted *longitudinal-longitudinal*, *transverse-longitudinal*, and *transverse-transverse* waves [11]. However, the approximation made by Einstein, to find the weak-field limit of the GWs, was considered debatable. In 1922 Arthur Eddington showed

that two of the wave solutions were artefacts of the choice of coordinate system [12], and threw the existence of the remaining transverse-transverse waves into doubt. The confusion over the existence of GWs continued, with one bizarre incident in 1936 serving to further muddy the waters. Einstein and Rosen submitted a paper entitled "Are there any gravitational waves?" to *Physical Review*, allegedly concluding that they do not. The anonymous referee of that paper (who would later transpire to have been the cosmologist Howard P Robertson) disagreed with the paper's conclusion (that GWs did *not* exist). Einstein's fury at the journal's editor for having sent the paper for peer review led to him withdrawing the publication, allowing him the time to be persuaded of the mistake which resulted in the work's erroneous conclusion.

The confused situation with the existence of GWs dogged the subject for several decades, with little progress made until Felix Pirani's work on describing GWs with respect to the Riemann tensor [13], a quantity which is, crucially, observable, and provides a coordinate-free description.

1.1.1 A wave solution to the field equations

The behaviour of gravitational fields in the presence of mass is modelled, in general relativity (GR), by the Einstein field equations,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{1.1}$$

Here $R_{\mu\nu}$ is the Ricci tensor, which represents the amount by which a sphere is distorted at a point in spacetime [2]; R the Ricci scalar [3], the

[2] More precisely, the Ricci tensor, which is the trace of the Riemann tensor, describes how the distance between the points within a volume varies as the entire volume is parallel-transported over a curved manifold, compared to the same movement over a flat manifold.

[3] The Ricci scalar is the trace of the Ricci tensor, and represents the deviation in the area of an $(N-1)$-dimensional sphere in a curved N-dimensional space compared to a flat N-dimensional space.

trace of the Ricci tensor; $g_{\mu\nu}$ is the metric tensor, which describes the local geometry of spacetime at any given point; and $T_{\mu\nu}$ the stress-energy tensor, which encapsulates the density and time-variation of energy and momentum at any given point in spacetime.

Contracting equation 1.1 with a timelike unit vector allows a reduction to a situation with a defined direction of time. This contraction leads to the revelation that spacetime curvature, R, is produced by the mass-energy density, since the contraction of a timelike vector with the stress-energy tensor returns simply the mass-energy density, ρ, in a dust of non-interacting masses. We then find

$$R = -4\pi G\rho,$$

which leads to the second-half of the famous aphorism,

> "Spacetime tells matter how to move; matter tells spacetime how to curve." —John Archibald Wheeler [14]

Despite their apparent simplicity the Einstein Field Equations are highly non-linear in the metric, $g_{\mu\nu}$, and only a few exact solutions are known. While some of these solutions, which include the Schwarzchild solution (for the behaviour of spacetime close to a non-spinning singularity), and the Kerr solution (for the spacetime close to a rotating singularity) are exact and tractable, the majority of scenarios lead to non-analytical solutions which must be explored numerically. While the Einstein equations provide this valuable insight into the behaviour of a static system, it is more often of interest to consider dynamic systems, and the time-evolution of the system. One approach to dealing with the non-linearity of the field equations is to "linearise" the theory, by assuming first a known solution of the field equations, $\eta_{\mu\nu}$, and then producing a small perturbation on that solution, $h_{\mu\nu}$. This approach is introduced in, for example, [4], and

provided that the size of the perturbation is sufficiently small, provides a means of investigating the model in "weak-field" scenarios. Such an approach is suitable for the study of the evolution of the weak-field. The metric in such a scenario takes the form

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \tag{1.2}$$

Allowing $\bar{h} \leftarrow h$, representing a rescaling [4] of the perturbation, $\bar{h} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$. We can then make a choice of gauge, the Lorentz (or Hilbert) gauge, by specifying $\bar{h}^{\mu\nu}{}_{,\nu} = 0$.

For convenience it is normal to work in "geometrised units", where the speed of light, $c = 1$. Doing so does, however, require care to include a c^{-1} factor in the value of x^0 when converting to natural units.

The derivative of the metric then describes the propagation of the perturbation,

$$\Box\,\bar{h}_{\mu\nu} \equiv \bar{h}_{\mu\nu,\alpha}{}^{\alpha} = 0, \tag{1.3}$$

where $\Box$ is the d'Alembertian box operator. This has plane-wave solutions of the form

$$\bar{h}_{\mu\nu} = \mathrm{Re}\left[A_{\mu\nu}\exp\left(\mathrm{i}k_\alpha x^\alpha\right)\right] \tag{1.4}$$

for a null wavevector, $\boldsymbol{k}$, orthonormal to an amplitude A. Provided a transverse, traceless[5] gauge is chosen, the amplitude tensor takes the form

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{xx} & A_{xy} & 0 \\ 0 & A_{xy} & -A_{xx} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.5}$$

It is clear from the form of the plane-wave solution in equation 1.4 that a wave propagates in spacetime in a manner quite different from a wave

[4] This rescaling of the metric has no physical consequence, but substanitally simplifies the number of quantities composing the Einstein tensor.

[5] The transverse-traceless gauge is convenient, since the metric perturbation is perpendicular to the wavevector in this gauge.

on, for example, the surface of a loch: rather than a vector perturbation, a GW propagates as a tensor perturbation. The conventional method used to visualise this effect involves considering the effect of a GW travelling perpendicular to a ring of test particles. For such a wave, propagating along the z-axis, in the transverse-traceless gauge, the only non-vanishing components of the strain are (returning to natural units by including c as a factor)

$$h_{xx} = -h_{yy} = \mathrm{Re}[A_{xx} \exp(-\mathrm{i}\omega(ct - z)],$$

and

$$h_{xy} = h_{yx} = \mathrm{Re}[A_{xy} \exp(-\mathrm{i}\omega(ct - z)].$$

The propagation can then be described as the superposition of two linearly-polarised components; the unit polarisation tensors can be derived from the coordinate basis as

$$\mathsf{e}_+ = \mathsf{e}_x \otimes \mathsf{e}_x - \mathsf{e}_y \otimes \mathsf{e}_y \tag{1.6}$$

$$\mathsf{e}_\times = \mathsf{e}_x \otimes \mathsf{e}_y + \mathsf{e}_y \otimes \mathsf{e}_x \tag{1.7}$$

As the wave passes orthogonally through the circular ring of test particles they will be distorted into an ellipse. For the $+$-polarisation the circle is stretched into an ellipse with semi-major axis first extending along the x-axis, relaxing back to a circle, and then again with a semi-major axis extending along the y-axis. This behaviour is depicted as a cartoon in figure 1.1.1.

Figure 1.1.1: The effect of a $+$-polarised GW on a circle of test particles as it propogates through the page (orthogonal to the ring). Time progresses horizontally along the x-axis from left to right.

Equivalently, the $\times$-polarisation produces a deformation rotated $45°$ relative to the $+$-polarisation; this is depicted in figure 1.1.2.

Figure 1.1.2: The effect of a $\times$-polarised GW on a circle of test particles as it propogates through the page (orthogonal to the ring). Time progresses horizontally along the x-axis from left to right.

The behaviour of "strong-field gravity", which is essential to understanding the production of GWs can only be practically probed using observed GWs from strong-field sources, such as binary black hole (BBH) coalescences. Systems such as these, which involve an accelerating mass, are capable of producing GWs according to the quadrupole formula, with the GW at a given time described by the three-dimensional tensor

$$h_{jk} = \frac{2G}{r}\frac{\mathrm{d}^2 Q_{jk}}{\mathrm{d}t^2} \tag{1.8}$$

where G is the gravitational constant, and Q_{jk}, the moment of inertia tensor, is defined as

$$Q_{jk} = \int \mathrm{d}^3 x \rho(\boldsymbol{x}) \left(x_i x_j - \frac{1}{3} r^2 \delta_{ij} \right) \tag{1.9}$$

for a mass density ρ, and coordinates x_j and x_k.

1.2 | Strain

The propagation of a GW will cause a relative displacement between test masses in spacetime. As a result, a GW will produce a relative strain, perturbing the normal metric. In the far-field approximation the metric, $g_{\mu\nu}$, can thus be described by

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu},$$

as first defined in equation 1.2, with the strain, $h_{\mu\nu}$ perturbing the underlying (potentially flat) metric $\eta_{\mu\nu}$.

The strain, as measured by a GW detector, can have two polarisation states, the $+$ state and the $\times$-polarisation state, which combine linearly, so that

$$h = ||\mathsf{A}_{+}h_{+} + \mathsf{A}_{\times}h_{\times}||, \tag{1.10}$$

with h_{+} being the amplitude of the strain in the A_{+} polarisation basis, and $h_{\times}$ the amplitude in the $\mathsf{A}_{\times}$ polarisation.

Finally, the overall measured strain in a detector will be the superposition of both the strain signal and noise (which is usually produced by movement of the detector's test masses due to effects other than spacetime perturbations). As such, the measured signal timeseries, $h(t)$, added to the noise timeseries $n(t)$ provides us with the total measured strain, $s(t)$,

$$s(t) = n(t) + h(t). \tag{1.11}$$

A similar quantity, the characteristic strain, is intended to account for integrating an inspiralling signal, leading to a straight-forward relationship between the characteristic strain and the signal-to-noise ratio (SNR).

Definition 1.2.1 (Characteristic strain)**.** The characteristic strain is a quantity which is intended to account for the effect of observing an inspiralling signal over the full period of the inspiral, thus integrating over a number of cycles of the signal. For a source with strain $\tilde{h}(f)$ as a function of frequency f it is defined as

$$[h_{\mathrm{c}}(f)]^2 = 4f^2 \left|\tilde{h}(f)\right|^2. \tag{1.12}$$

If we consider only the noise component of the recorded data, $n(t)$ from equation 1.11, then we can define the (one-sided) power spectral density (PSD) of the noise, $S_n(f)$, such that

$$\langle \tilde{n}(f)\tilde{n}^*(f)\rangle = \frac{1}{2}\delta(f - f')S_{\mathrm{n}}(f) \tag{1.13}$$

where $\tilde{n}(f)$ is the Fourier transform of the time-domain noise measurement $n(t)$, f is the frequency, and δ is the Kronecker delta function. The angle bracket notation $\langle \tilde{n}(f)\tilde{n}^*(f) \rangle$ represents an average over many instances of the noise power, which is required in order to achieve a good estimate of the PSD. This representation of the noise makes the assumption that it is stationary. In reality, this is not strictly true, but they are reasonable approximations for many applications; non-stationarities in the noise can become a problem for GW detection algorithms however, and these are discussed in more detail in section 1.4.2.

In analogy to the characteristic strain from definition 1.2.1, we can define the *characteristic noise*:

$$[h_{\mathrm{n}}(f)]^2 = f S_{\mathrm{n}}(f). \tag{1.14}$$

As noted by [15] this allows the integration of the strain compared to the noise budget of a given detector to be estimated "by eye", when displayed on a log-log plot.

The ability to detect a signal in a noisy data stream requires that the signal has sufficient power to be distinguished from the underlying noise. In GW analysis it is normal to express the strength of such a signal by reference to its SNR. This is defined with reference to the optimum filter for the signal, which is the Weiner filter (see [15] for a discussion of this). This filter gives an expression for the SNR, ρ [6], in terms of the signal strain in the frequency-domain, $\tilde{h}(f)$, and the noise PSD, $S_{\mathrm{n}}(f)$:

$$\rho^2 = \int_0^\infty 4\frac{|\tilde{h}(f)|^2}{S_{\mathrm{n}}(f)}\,\mathrm{d}f = \int_{-\infty}^{\infty} \left[\frac{h_{\mathrm{c}}(f)}{h_{\mathrm{n}}(f)}\right]^2 \mathrm{d}(\log f). \tag{1.15}$$

[6] Note here that ρ is routinely used to represent an SNR in signal processing, but this does introduce a confusing multiplicity, given the frequent use of ρ for the mass density in physics.

1.3 | Detecting gravitational waves

Despite Pirani's work simplifying the description of GWs in GR, it would take until 1957 for his arguments to gain prominence. The Chapel Hill Conference of 1957 brought together around 40 physicists at the University of North Carolina, Chapel Hill, with discussions focussed around gravitation and GR [16]. It was during a session of this meeting chaired by Hermann Bondi that Richard Feynman is credited with developing the "sticky bead" argument. Feynman used Pirani's formulation to argue that a device could be constructed which would measure the energy carried by a GW.

Consider two beads on rigid rod, which are free to slide along the rod, experiencing some friction. As a GW moves along the rod the length of the rod will remain fixed thanks to inter-atomic forces, but the proper distance between the two beads will change. This will result in the beads rubbing on the rod, generating friction, and thus heat, which can be measured [17].

One of the attendees of the meeting was Joseph Weber. Weber was the first person to propose a practical GW detector [18] while at the University of Maryland. He later went on to construct a resonant bar detector (see section 1.3.1) from which he claimed the first detection of signals originating in the centre of the Galaxy, in 1969 [19, 20, 21].

Numerous attempts to confirm his findings were unsuccessful, including searches in Ronald Drever's group at the University of Glasgow [22] in the United Kingdom; at Bell Labs [23, 24, 25] in the United States; at Munich [26] in Germany; at Moscow [27] in Russia; and at Tokyo [28] in Japan. While Weber's original detections were soundly refuted by the community there is little doubt that the announcement led to a flurry of activity in the field. This ultimately lead to the development of modern cryogenic resonant bars, such as ALTAIR [29], ALLEGRO [30], NAU-

TILUS [31], and EXPLORER [32]; and laser interferometers.

Laser interferometers, of which advanced LIGO is an implementation, were the result of a quest for both higher sensitivities and greater bandwidth. The possibility of using a Michelson interferometer to measure the distance between test masses in order to detect gravitational radiation originated in Moscow[33] in 1963, and again in 1966 [34].

Robert Forward, a former student of Weber, who had been involved in the construction of the original Weber Bar, was the first to work on the development of an interferometric detector, at Hughes Research Laboratory in the early 1970s, with the development of a "laser transducer" [35] in 1971. This lead to the development of an 8.5-metre detector [36], which failed to show any signal correlation with the bar detectors at Argonne, Glasgow, Friscati, or Maryland.

This approach was followed early-on by Scottish and German groups as a means of improving on resonant bar sensitivities, with a 3-meter and later a 30-meter prototype detector constructed at Garching in the late 1970s [37, 38] which used optical delay lines, and a 1-meter prototype, and later a 10-meter instrument was built at Glasgow in the early 1980s [39, 40], which used Fabry-Perot cavities. The Glasgow detector was the spiritual predecessor to the CalTech 40-meter prototype [41].

The increasing maturity of technology developed by these prototypes lead to the construction of the first generation of long-baseline detectors. The group at Glasgow had aspirations to construct such a detector in Scotland [42], while the group in Garching had similar plans for a German detector. While neither detector came to fruition, a smaller-scale, joint German-UK detector, GEO600 [43] was constructed near Hannover. The TAMA detector was built in Tokyo [44]. These would be joined by the three kilometre-scale joint Caltech-MIT initial LIGO detectors [45], located at two sites in the USA, and the joint Italian-French detector

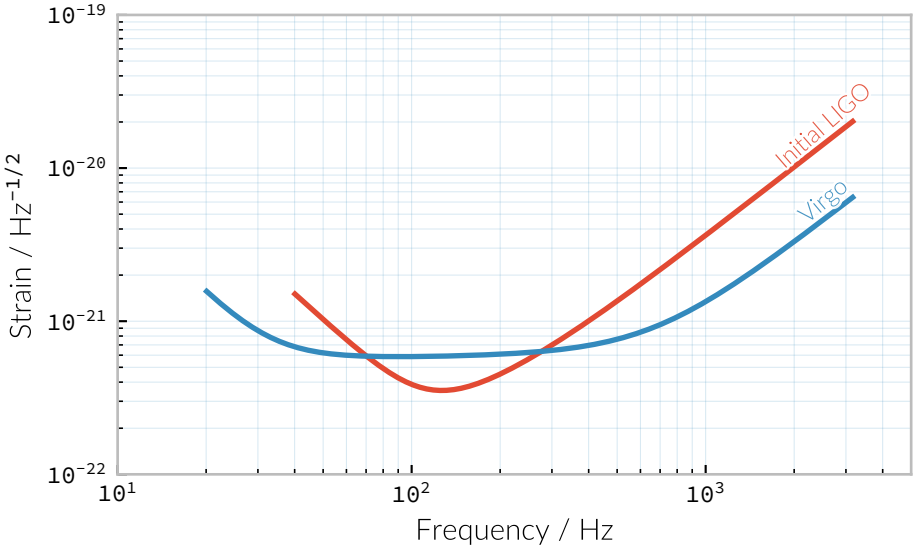


Figure 1.3.1: The approximate ASDs for the first generation of large-scale interferometers: initial LIGO (red), and Virgo (blue), derived from the fits in table 1 of [52].

Virgo [46], near Cascina. These detectors were operated during the 2000s, and while none of them made a detection of GWs, they provided valuable astrophysical results by placing astrophysical limits on the strength of the stochastic GW background [47], production of GWs by pulsars [48] and gamma ray bursts [49], and the rate of compact binary coalescence in the local universe [50, 51].

Figure 1.3.1 is a plot of the noise amplitude spectral density (ASD) of the first generation of interferometric detectors, which demonstrates the wide range of frequencies which detectors of this type are capable of measuring GW strain over.

The initial-generation of detectors were upgraded during the first half of the 2010s, leading to Advanced LIGO [54] which resumed observations in September 2015, with the Advanced Virgo detector [55] joining in summer 2017 to conduct joint observations with its counterparts in the USA. The GEO600 detector was the first of the initial detectors to be fully upgraded

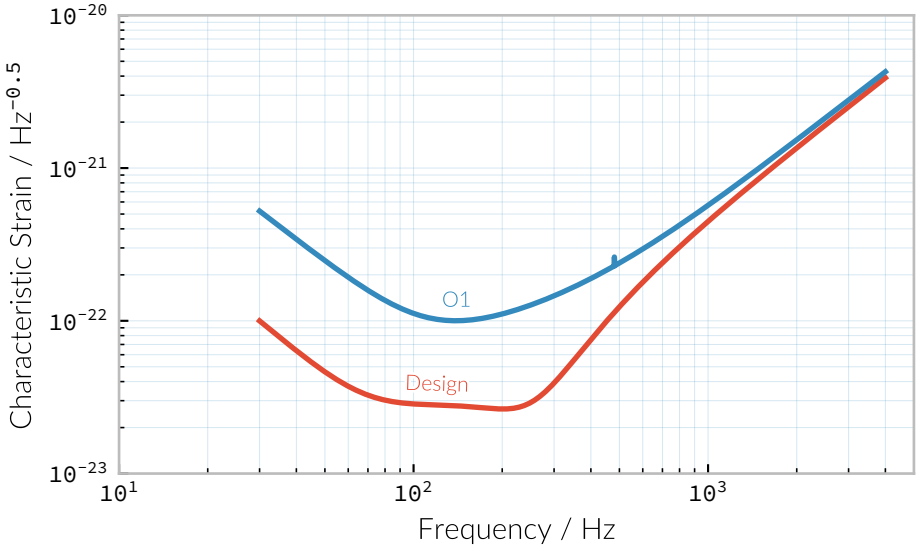


Figure 1.3.2: The predicted ASD of the Advanced LIGO detectors within their sensitive band, at design sensitivity (from the fit in table 1 of [52]), relative to the estimated sensitivity of the two interferometers in their first observing run (O1) [53].

as part of the GEO-HF project [56], with improved sensitivity at high frequencies. Japanese efforts have focused on the development of Kagra (formerly Large scale cryogenic gravitational wave telescope (LCGT)), a cryogenic interferometer located deep underground in the Kamioka mine [57], which is expected to join the third observing run of advanced LIGO. The construction of a third LIGO interferometer in India using the mothballed second detector from the Washington site has now moved into its initial stages, with the prospect of this detector joining the network by the mid-2020s. Figure 1.3.2 depicts the anticipated ASD of the advanced LIGO detectors once they have reached their design sensitivity, which is expected within the next five years.

The second-generation detectors, specifically the two advanced LIGO detectors responsible for the first discovery of GWs [58], have successfully demonstrated the ability of interferometry to observe the gravitational

universe. This said, future improvements in sensitivity are highly desirable, but are likely to be even more technically challenging than the transition from resonant bars to laser interferometers.

In order to improve the bandwidth of detectors a location of minimal *Newtonian noise* (see 1.4.1.3), which results from variation in the local gravitational field, must be found, which ultimately mandates the placement of an interferometer in space. The earliest proposals for a space-based detector came in the form of LAGOS, which originated as a concept at the University of Colorado under Jim Faller and Peter Bender [59]. These proposals would develop into Laser Interferometer Space Antenna (LISA) [60], which is likely to launch in the 2030s. The technology demonstration mission for LISA, *LISA Pathfinder* was launched in December 2015, and its main mission was completed successfully in early 2016 [61]. The LISA detector will be sensitive in the milli-hertz region of the GW spectrum, and will be capable of observing binary inspirals at a much earlier stage in their evolution than the advanced ground-based detectors, as well as the galactic population of low-mass binaries, such as binary white dwarfs. A Japanese proposal, Deci-hertz Interferometer Gravitational wave Observatory (DECIGO) [62], would observe in the decihertz regime using a complex arrangement of six spacecraft in a star-of-David configuration.

There are also plans for more sensitive detectors on the ground. The Einstein telescope is a European proposal for an underground kilometre-scale detector in a triangular configuration, using a "xylophone" configuration to improve broadband sensitivity compared to the second-generation of detectors; its scientific aims include providing more sensitive tests of GR than are possible with the advanced detectors [63]. The prospect also exists for larger surface-based detectors, such as Cosmic Explorer, which would have an arm-length of 40-km [64], initially using technology currently under development for the upgrade of advanced LIGO, but later

incorporating cryogenic technology, such as those under development for Kagra [65]. There are also proposals for upgrades of the advanced detectors to use squeezed light to reduce quantum noise [66], the use of speedmeters [67, 68], or atom interferometry [69, 70, 71].

At the very low-frequency limit of the GW spectrum the bulk of detection efforts are based around pulsar timing arrays, which promise the detection of GWs by precision measurements of pulse arrival times from a number of pulsars distributed across the sky. By observing correlated delays [72] in arrival times the presence of a very long wavelength GW can be inferred. There are a number of collaborations actively producing pulsar observations with the aim of detecting GWs: the European Pulsar Timing Array (EPTA) [73], NANOGrav [74], the Parkes Pulsar Timing Array (PPTA) [75], and the International Pulsar Timing Array (IPTA) collaboration [76].

1.3.1 Resonant bar detectors

The original GW detectors developed by Weber in the 1960s were an early example of a category of detector now known as a *resonant bar*. These detectors work on the principal that variations in the Riemann tensor will drive oscillations between two masses. If the Riemann tensor inside a crystal varies, the stress tensor of the crystal will also vary, and if the crystal is piezoelectric, this will in turn produce a change in the polarisation in the material. In Weber's earliest design [18] the change in the electric field in a piezoelectric crystal would be monitored through changes in the voltage across the crystal with a low-noise radio receiver. Such an arrangement relied on a single instrument; the rotation of the Earth would produce a variation in the strength of what was expected to be a continuous GW signal measured by the instrument, allowing its direction to be determined.

Alternatively Weber proposed an arrangement of two instruments with cross-correlated outputs which he imagined would remove the need for diurnal variation in this process. A major complication of this approach was the need to have low-noise amplification of the measured electric field from the crystal, which Weber had hoped (in 1960) would be realised through the use of masers. By 1966 Weber's detector, which consisted of an aluminium bar weighing approximately $\sim 1360\,\mathrm{kg}$, fitted with quartz piezoelectric strain gauges, was capable of making strain measurements around $h \sim 10^{-16}$, with the pre-amplifier cooled with liquid-helium.

The 1990s brought a second generation of resonant detector design, and an international network of five detectors, which were cooled to cryogenic temperatures to reduce thermal Nyquist noise within the bar. A mechanical resonator, which was tuned to a specific frequency was then attached elastically to one face of the bar. The displacement between this resonator and the bar face was measured via the capacitance between the bar face and the secondary resonator. The cryogenic generation of detectors were capable of reducing the noise strain in the detector to around $10^{-22}\,\mathrm{Hz}^{-1/2}$.

While the sensitivity of bar detectors was much improved over three decades of development, the narrow bandwidth (around $1\,\mathrm{Hz}$ centred around the resonance frequency of the detector) substantially reduced the quantity of the GW signal which can be measured from most plausible astrophysical sources. This has caused resonant bar technology to struggle to compete with detectors based around laser interferometry (see section 1.3.2) which typically have bandwidths on the order of $10^3\,\mathrm{Hz}$.

Despite this, development of resonant mass antennas is ongoing. In addition to both NAUTILUS and AURIGA, there are two spherical cryogenic detectors, MiniGRAIL [77], and Mario Schenberg [78], which hope to be able to make GW measurements at higher frequencies than the

current generation of interferometric detectors through cooling to $50\,\mu\mathrm{K}$.

1.3.2 Interferometric detectors

Gravitational-wave detectors which use beams of light, such as interferometers and pulsar timing arrays rely on measuring the the travel time of a beam of electromagnetic radiation between two points, and the effect that a GW has on this time. A full treatment of this is given in [52], but in summary, if a GW is not present within a detector, the travel time of a beam in the detector will be constant. A beam of light is generated at a proper time t, and is received by a sensor at a proper time τ. With no GW the proper distance between the two clocks is L. If the beam of light is generated with some sort of time-stamp, then the receiving sensor can measure the time of arrival of these time-stamps. If no GW is present the rate will be constant, and we can choose a unit of time in which this rate is unity.

If a GW is introduced, which produces a strain, $h_+(t)$, in the plane of the beam, the change in the arrival time of the beam will be changed. If a beam leaves the transmitter at time t, when the GW strain will be $h(t)$, it is received at a time τ, when the GW strain will be $h(t + (1 - \cos(\theta))L)$, with θ the angle between the direction of the beam and the direction of GW propagation. This means that the arrival rate is changed compared to the emission rate by

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = 1 + \frac{1}{2}(1 + \cos\theta)\left\{h_+\left(t + [1 - \cos\theta]L\right) - h_+(\tau)\right\}. \tag{1.16}$$

By arranging the detector to reflect the beam back to the originating clock, it is possible to measure the round-trip time using only one clock. In this arrangement we must account for the GW having a different strength

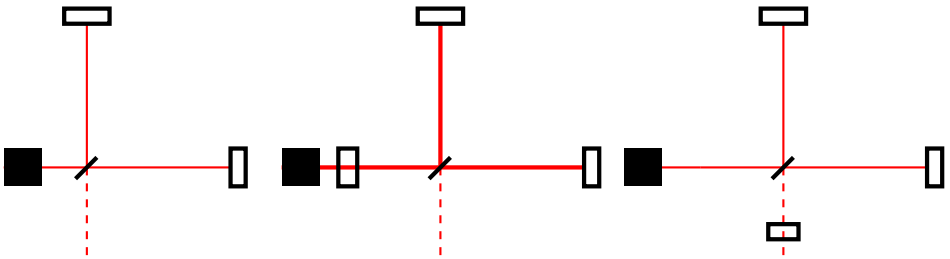


Figure 1.3.3: **Left**: A simple Michelson interferometer, composed of a light source (black box), a beam splitter (heavy black line), and two end mirrors (white boxes). **Centre**: A Michelson interferometer with an additional power recycling mirror, placed between the beam source and the beam splitter. **Right**: A Michelson interferometer with a signal recycling mirror, placed between the beam splitter and the output port.

one the return trip, and so equation 1.16 becomes

$$\frac{\mathrm{d}t_{\mathrm{round}}}{\mathrm{d}t} = 1 + \frac{1}{2}\Big[(1-\cos\theta)h_+(t+2L) - (1+\cos\theta)h_+(t) + 2\cos\theta\; h_+\left(t + L[1-\cos\theta]\right)\Big], \tag{1.17}$$

which is often called the *three-term* relation.

1.3.2.1 Operation of a Michelson interferometer

A Michelson interferometer is an optical device which is capable of measuring the difference in length between two optical paths to sub-wavelength precision. A Michelson interferometer can be constructed using a beam splitter and two mirrors, in the configuration presented in the left panel of figure 1.3.3. The input beam is split along the x and y directions, and reflected back to the beam splitter. At the beam splitter the two beams will interfere: in the standard Michelson setup this will result in constructive interference if the arms have identical lengths, and a beam will be produced at the output (the dashed red line). If the arms' relative lengths change a pattern of interference fringes will be visible at the output of the interferometer.

This means that we can consider an interferometer with two arms to consist of one arm which acts as the time standard, against which the variations of the other can be measured. However, such an arrangement also means that if the effect of a GW is the same on both arms it will not be detectable, and will be most detectable if one arm is extended while the other is contracted by the same amount.

1.3.2.2 Power recycling

The optimal signal-to-noise ratio can be achieved from an interferometer when the arm lengths are configured so that when no GW is present in the interferometer the interferometer beams interfere destructively [79]. If the mirrors absorb little energy, the light will then be reflected back towards the laser, and by placing a mirror between the laser and the beam splitter a resonant cavity can be formed (see the middle panel of figure 1.3.3), allowing the power in the interferometer to build up. This allows a less powerful laser to be used as the input for the interferometer, with a laser capable of providing several kilowatts of power inside the interferometer [80].

1.3.2.3 Signal recycling

Signal recycling can be used to tune the bandwidth of an interferometer, and to increase its sensitivity by re-injecting the interferometer's output signal to the interferometer, achieving resonance, which increases the signal-to-noise ratio of the signal. This is possible thanks to the sidebands on the beam which are produced by the GW not interfering destructively.

To perform signal recycling a mirror is added between the beam splitter and the readout port of the interferometer [81, 82], with this configuration illustrated in the right panel of figure 1.3.3.

1.3.2.4 Fabry-Perot cavities

For a ground-based interferometer, which has an arm-length of 4-kilometres, the light travel time within the arm is of the order $10^{-5}\,\mathrm{s}$. The period of a GW which the detector is sensitive to, around $10^{-2}\,\mathrm{s}$, is much greater than this travel time [83]. As a result it is advantageous to allow the beam to remain within the arm for longer than one round-trip. By setting the arm up as a cavity the effective length of the arm can be increased; a finesse of 100 will then increase the effective length of the arm 100-fold. This in turn increases the apparent change in the arm length by a factor of 100, and substantially aids the sensitivity of the detector.

In Advanced LIGO, for example, the main arms form a Fabry-Perot cavity, with a finesse of 450 [54]. This is formed by placing a mirror between the beam-splitter and the end mirror in each arm.

1.3.2.5 Antenna response of the detector

The arrangement described in section 1.3.2.1, whereby one arm is used as the timing reference causes the detector to be incapable of detecting signals if both arms are affected equally by a GW. The angle between the propagation of the GW and the detector (in addition to the polarisation of the GW) will determine the effect on each arm. This results in an interferometric detector having a varying sensitivity to sources across the sky, which is conventionally treated as an antenna pattern, in analogy to the similar concept in radio astronomy. For a GW approaching the detector from an azimuth (relative to one of the arms) and altitude (relative to the plane of the detector), (α, δ) on the sky these patterns for the +-

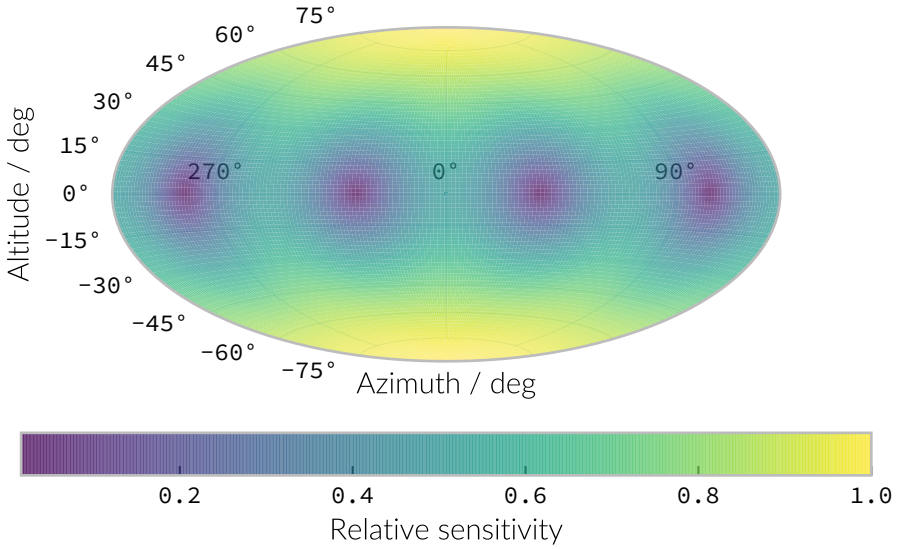


Figure 1.3.4: The normalised antenna pattern, in response to +-polarised GWs, of a signle two-armed interferometric detector with a $90°$ arm separation, with axes in the x-y plane. Here the azimuth positions assume that one of the arms is oriented north-to-south (along the y-axis) and the other east-to-west (along the x-axis); an appropriate rotation should be added to account for alternative orientations.

and $\times$-polarisations, F_+ and $F_\times$, will be

$$F_+ = \frac{1}{2}(1 + \sin^2\delta)\cos 2\alpha \cos 2\psi - \sin\delta \sin 2\alpha \sin 2\psi \tag{1.18}$$

$$F_\times = \frac{1}{2}(1 + \sin^2\delta)\cos 2\phi \sin 2\psi - \sin\delta \sin 2\phi \cos 2\psi \tag{1.19}$$

for ψ the polarisation angle of the GW, which corresponds to the rotation of the basis vectors defining the polarisations of the GW compared to the detector [52]. The +-polarised response is plotted in figure 1.3.4, which clearly depicts the four regions of low sensitivity.

The overall measured strain, $h(t)$ in a detector from a GW with components $(h_+, h_\times)$ will then be

$$h(t) = F_+(t)h_+(t) + F_\times(t)h_\times(t) \tag{1.20}$$

While this antenna pattern has the effect of reducing the sensitivity of

the detector to some areas of the sky, it provides additional information relating to the direction of the GW. This information can be utilised if a network of detectors is available, as if a signal is detected in similar detectors located elsewhere, but not (or barely) detected by another, it may be possible to infer that the signal originated in the direction of the one of the non-detecting detector's "blind spots". Such an inference was valuable in the localisation of the source of GW170817 [84], which had a noticeably weak signal in the Virgo detector.

1.3.2.6 Localising a gravitational wave signal

If a network of at least two geographically separated detectors observes a signal it is possible to ascertain the location in the sky, $\hat{\mathbf{\Omega}}$, from the difference in arrival times between the two sites. For a detector at a position, $\boldsymbol{r}_D$, and an arbitrary reference location, $\boldsymbol{r}_0$, this time delay, δt, will be

$$\delta t(\hat{\mathbf{\Omega}}) = \frac{1}{c}(\boldsymbol{r}_0 - \boldsymbol{r}_D) \cdot \hat{\mathbf{\Omega}} \tag{1.21}$$

This allows the location of the signal to be confined to a ring on the sky corresponding to constant Δt. Timing uncertainty in the signal, which arises both from clock uncertainties and uncertainties in defining a reference point in the received signal increase the area of this region. As more detectors are added to the network it is possible to reduce this area, as increasing the number of detector pairs works to reduce the sky area compatible with the observed delay times.

Additional localisation information can be attained from the observed amplitude of the signal in each detector. The signal will be convolved with the antenna pattern (see section 1.3.2.5); as each detector is insensitive to some regions of the sky, the total plausible localisation of the signal is reduced.

1.3.3 Ground-based interferometers

While there are attractions to being able to place an interferometric GW observatory in space, practical concerns have so-far constrained these detectors to being placed on the ground (or, in the case of Kagra, under it). Fortunately, a considerable amount of science is possible with ground-based detectors, within the acoustic band of frequencies (above around 10-hertz). As a result considerable effort has been put into the development of detectors which can overcome the noisy environment which these detectors experience, which has so-far culminated in the construction of the advanced LIGO observatories, and the advanced Virgo observatory. In the near future these are likely to be joined by Kagra and an additional LIGO detector in India.

Future developments in ground-based interferometry are likely to force the detectors underground in order to mitigate seismic and Newtonian noise (see section 1.4); Kagra has already been located in a mine, while a plan for a future subterranean detector is the Einstein Telescope.

1.3.3.1 Advanced LIGO

The Advanced LIGO detectors are considered second-generation interferometric GW detectors, located at two observatories in the United States of America. LIGO Livingston Observatory (LLO) is located in woodland outside the town of Livinston in Louisiana, while LIGO Hanford Observatory (LHO) is located on the Hanford Reservation in the State of Washington.

The advanced LIGO detectors replaced the first-generation Initial LIGO detectors, and share the same facilities as their predecessors[7], and

[7] With the exception of the 2-kilometre detector at the LHO site, which was not upgraded; the unusued infrastructure from this detector is earmarked for a future LIGO detector in India.

Parameter	Value
Arm length	3994.5 m
Arm finesse	450
Laser wavelength	1064 nm
Input power	125 W
Test-mass mass	40 kg

Table 1.3.1: The basic parameters of the advanced LIGO detectors, from [54].

like them are 4-kilometre long interferometers with a Fabry-Perot cavity in each arm, with a finesse of 450. The detectors improve their sensitivity compared to the initial generation detectors through the use of signal recycling, a technology pioneered in the GEO600 detector, and have quadruple mirror suspensions which use fused silica fibres to provide seismic islolation [85, 86]. Combined, the improvements to the design of the detectors allowed a ten-fold improvement in sensitivity in the most sensitive frequency region (around 100 Hz) compared to the initial LIGO detectors, as can be seen in the difference between the sensitivity curves in figures fig:detectors:interferometers:firstgen and fig:detectors:aligo-asd.

The first continuous observations with the advanced detectors started in September 2015. During the first observing run[8] the detectors made three detections of coalescing BBH.

1.3.3.2 Advanced Virgo

Similarly to advanced LIGO, the advanced Virgo detector is a second-generation interferometric detector which replaced a first-generation detector. Located in Cascina, Italy, this detector has a number of design choices

[8]The standard nomenclature for advanced-era observing runs is of the form "O<number>", so the first observing run was "O1". These are independent of the actual detectors involved in the run, so when advanced Virgo started observations concurrently with the advanced LIGO detectors during its second observing run, the run was known universally as "O2".

which are distinct compared to the LIGO detectors, choosing, for example to use "super attenuators" rather than the quadruple suspension system of LIGO to provide seismic isolation. Additionally, the detector's arm cavities are shorter than those of advanced LIGO, extending 3-kilometres compared to LIGO's four.

1.3.3.3 Kagra

The final "advanced era" detector design which is under development is that of Kagra (previously known under the moniker LCGT) [87]. Kagra has claim to bridge the technological divide between the second and third generation of GW detectors, as it is expected to be the first interferometric detector to employ cryogenic technology. The use of cryogenically-cooled mirrors is designed to reduce thermal noise originating in the mirror coatings (see 1.4.1.1), but presents a number of technological challenges which ambient-temperature detectors avoid. Additionally, in contrast to LIGO and Virgo, Kagra will be located underground (in a disused part of the Kamioka mine complex). This principle is expected to be used for the Einstein Telescope, and reduces the impact of some forms of Newtonian noise (see section 1.4.1.3) on the detector, and thus improves its low-frequency sensitivity. Unlike planned third-generation detectors, however, Kagra will have an arm length of 3-km, around an order of magnitude smaller than future subterranean detectors are anticipated to be.

1.3.3.4 Einstein Telescope and Cosmic Explorer

The two plans for third-generation detectors which are currently under consideration are Cosmic Explorer, which is likely to be located in the USA, and Einstein Telescope, likely to be located in Europe. A number of technological advances are anticipated which will allow a considerable

increase in sensitivity over the current generation of detectors, in addition to increased arm cavity lengths (40-kilometres in the case of Cosmic Explorer, and 30-kilometres for Einstein Telescope). The sensitivity improvements in this generation of detectors should allow the detection of compact binary coalescence (CBC) events to very high ($z > 10$) redshifts at high SNR [88]. In addition to having longer arm cavities than current detectors, Einstein Telescope will be placed underground in an attempt to mitigate Newtonian noise (see section 1.4.1.3).

1.3.4 Space-based interferometers

While ground-based interferometers have the advantage of accessibility, and consequently fairly affordable construction costs, great advantage is to be had in placing an interferometer in space. Some noise sources which detectors such as LIGO must contend with, such as seismic noise, are completely absent, and greater freedom is afforded in the size of the interferometer, with the absence of a need to purchase and prepare land for the observatory. In exchange for these advantages space-based interferometers present a number of technological hurdles, such as maintaining sufficiently stable orbital configuration to allow interferometry to be carried-out, and reduced sensitivity, as constructing a Fabry-Perot cavity in the comparatively poor vacuum around the L1 point is not feasible.

Despite these difficulties, space-based detectors represent the majority of feasible concepts for detectors sensitive to low frequency emission. The following sections contain further details of the LISA and DECIGO mission proposals, but numerous other proposals for space-based detectors exist, including Geostationary LISA (gLISA) [89, 90], which proposes using off-the-shelf satellites to form a detector constellation in geostationary (rather than heliocentric) orbit. The TianQin mission proposal [91] also

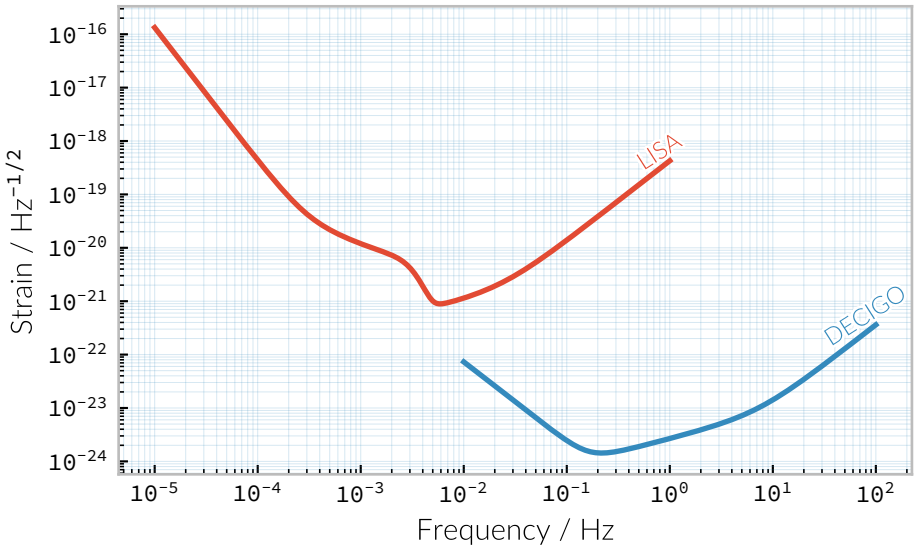


Figure 1.3.5: The ASD of the LISA and DECIGO detectors within their sensitive band, at design sensitivity. The curve for LISA is based on the prediction outlined in [92], while the DECIGO curve is based on the approach in [93].

uses such a technique, with the aim to have a shorter development time than rival concepts such as LISA.

1.3.4.1 Laser Interferometer Space Antenna

LISA is a planned space-based GW observatory, under development by the European Space Agency, which would be placed in a heliocentric orbit at the L1 Lagrange point. In comparison to the kilometre-scale arms of second-generation ground-based detectors such as LIGO, LISA is proposed to have arms which are 2.5 million kilometres long, giving the detector much greater sensitivity at low frequencies than is possible with ground-based detectors. The ASD of LISA is plotted in figure 1.3.5.

The LISA mission was preceeded by LISA Pathfinder, a technology demonstration mission, launched in December 2015.

1.3.4.2 DECIGO

DECIGO [62] is a proposed space-based GW observatory which is designed to observe the deci-hertz GW regime. Ground-based detectors are sensitive to frequencies above around 10 Hz, and the LISA mission is designed to observe frequencies below 1 Hz. This leaves a region which is unobserved, centred approximately around 10 Hz, which overlaps with less sensitive regions of the LISA and ground-based detectors passbands. The ASD of DECIGO is plotted in figure 1.3.5.

A DECIGO cluster will consist of three spacecraft in a triangular configuration, forming three Fabry-Perot cavity cavities with lengths around 1000 km. Four of these clusters, placed in heliocentric orbits, will form the entire observatory constellation, with two of the clusters arranged in a nearly-overlapping "Star-of-David" geometrical configuration [94].

1.3.5 Pulsar timing

Pulsar timing relies on observations made of the arrival times of pulses from millisecond pulsars. In comparison to an interferometer, where the measurement of the detector's arm is made by observing the phase of the laser beam over a scale of a few kilometres (in the case of a ground-based detector such as LIGO), or even a few gigametres (in the case of LISA), pulsar timing arrays provide an arm length on the scale of parsecs. Accordingly, they are sensitive to much lower frequencies than man-made detectors.

If a pulsar is treated as a clock which produces pulses at predictable intervals, any discrepancy between the predicted arrival time and the observed arrival time may be attributed to some effect along the line of sight. The phase, ϕ, of the signal from a pulsar which has a rotation frequency and phase at a time, t_0, of respectively ν_0 and ϕ_0, and a spin-

down rate, $\dot{\nu}$, can be found as

$$\phi = \phi_0 + \nu_0(t - t_0) + \frac{1}{2}\dot{\nu}(t - t_0)^2, \tag{1.22}$$

at time t. By setting the observational epoch to begin with the first observation (so that $t_0 = 0$), the time of arrival, t of the N-th can be related as

$$N = \nu_0 t + \frac{1}{2}\dot{\nu}t^2 + \epsilon, \tag{1.23}$$

for ϵ a noise term which results from any effects along the line of sight.

The effect of a GW on the arrival time of a specific phase can be found from equation 1.16; the presence of a GW along the line of sight between the pulsar and the observer (conventionally located at solar system barycentre to remove various timing effects related to the movement of the Earth in the solar system) will be seen in the amplitude of the ϵ term of equation 1.23. GWs are not the only potential source of additional "timing noise" however, as any variation in the gravitational field in the vicinity of either the pulsar or the observer will contribute to variation in ϵ. In order to detect GWs it is therefore necessary to observe a number of pulsars, and compare correlations in the ϵ data (known as "timing residuals") for each of them.

The correlation between pulsars is dependent upon their angular separation, ζ, in the sky [72], and given by the "Hellings-Downs curve", which provides the sky- and polarisation-averaged response of a pair of pulsar lines-of-sight to a plane GW, and has analytical form

$$\chi(\zeta) = \frac{1}{2} - \frac{1}{4}\left(\frac{1 - \cos\zeta}{2}\right) + \frac{3}{2}\left(\frac{1 - \cos\zeta}{2}\right)\log\left(\frac{1 - \cos\zeta}{2}\right), \tag{1.24}$$

for ζ the angular separation of the Earth-pulsar baselines for each pulsar. This relationship is plotted in figure 1.3.6.

In the case of a pulsar timing array there will be numerous pulsars; the Hellings-Downs correlations for each can be calculated as a pairwise

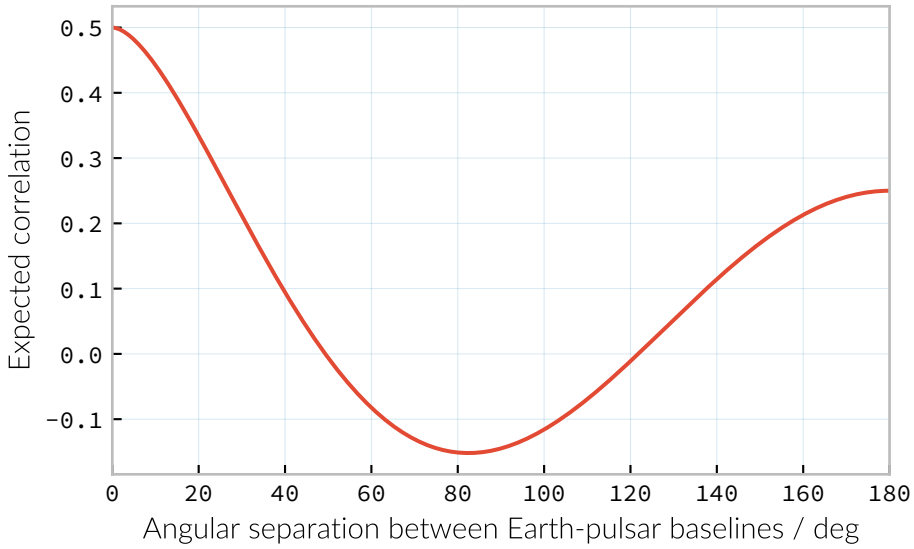


Figure 1.3.6: The Hellings and Downs curve giving the expected correlation between a pair of Earth-pulsar baselines with a given angular separation.

matrix, $\chi_{ij} = \chi(\zeta_{ij})$ for ζ_{ij} the angular separation between pulsars i and j within the array of M pulsars, with $i, j \in 1, ..., M$. These correlations, along with the timing noise of each pulsar, can be used to construct the PSD of the array.

1.3.6 Other approaches

A number of other techniques have been used to place limits on various forms of GW emission, including Doppler ranging of spacecraft [95], astrometry using GAIA observations [96], the measurement of the Earth's normal modes [97]. Proposals for alternatives to light-based interferometry also exist in the form of atom interferometers [98, 99].

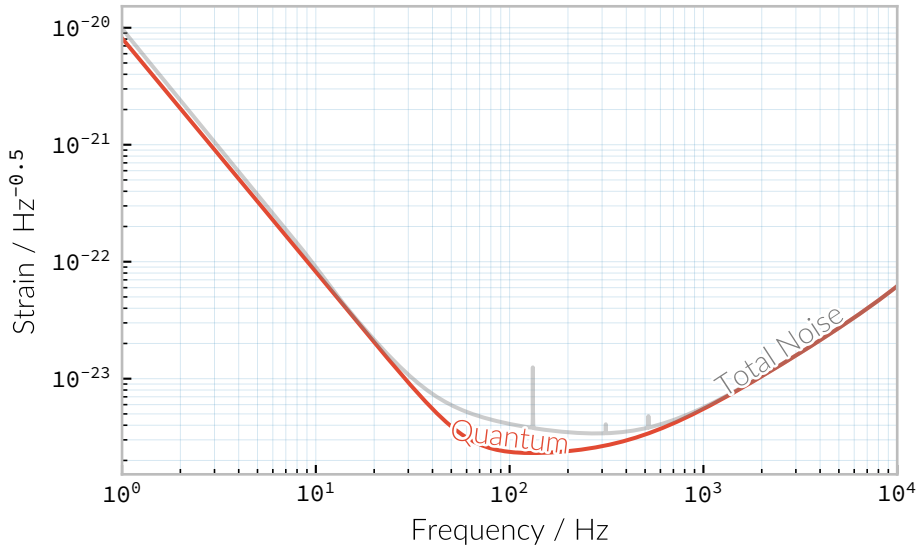


Figure 1.4.1: The contribution to the advanced LIGO ASD from quantum noise. These curves were calculated using the `pygwinc` library [100].

1.4 | Noise sources

Given the small strain amplitudes of GWs, and the correspondingly small displacements they produce in a detector, the detector data is normally dominated by noise. This noise limits the range over which a detector is sensitive to GWs, so understanding the sources of noise, and mitigating them is the most effective means of improving their sensitivity to astrophysical sources.

Noise sources are split broadly into two categories: instrumental sources, and facilities source. The former includes noise sources which are due to the equipment used to construct the detector, the latter are a result of physical properties of the observatory's site and infrastructure.

1.4.1 Quantum noise

One of the major sources of instrumental noise in detectors such as advanced LIGO is from quantum fluctuations in the intensity of the photon field in the detector arms. This manifests itself through two processes. The first is as radiation pressure noise; a change in the photon flux reflecting off the mirror will lead to a fluctuation in the radiation pressure exerted on the mirror (and hence the test mass). The PSD of this noise, given a power P circulating in the arm cavities, with a wavelength λ, and with the mass of the test mass m is

$$S(f) = \frac{1}{mf^2L}\sqrt{\frac{\hbar P}{2\pi^3 c\lambda}}, \tag{1.25}$$

at a given frequency f (with $\hbar$ the reduced Planck constant), for a detector with arm-length L [80]. Radiation pressure can be mitigated by increasing the power circulating in the arms, however this must be balanced against the increased shot noise introduced by the increased power.

Shot noise results from quantum fluctuations in the photodiode which measures the output signal from the interferometer. For the same interferometer properties listed for the radiation pressure noise in equation 1.25 this is

$$S(f) = \frac{1}{L}\sqrt{\frac{\hbar c\lambda}{2\pi P}}. \tag{1.26}$$

As a result increasing the laser power will increase the shot noise at high frequencies.

The combined quantum noise for advanced LIGO is shown, alongside the total noise budget of the detector in figure 1.4.1.

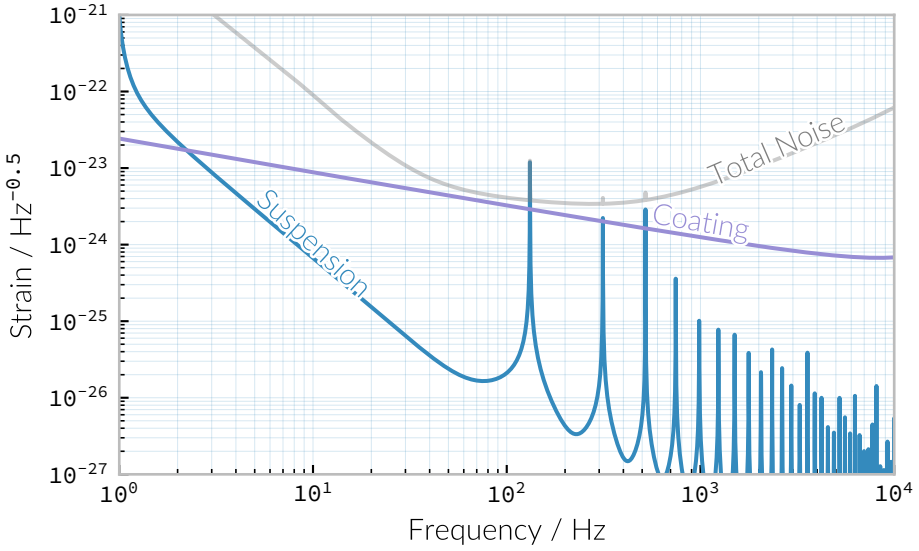


Figure 1.4.2: The contribution to the advanced LIGO PSD from thermal noise. These curves were calculated using the `pygwinc` library [100].

1.4.1.1 Thermal noise

Thermal noise primarily affects the low-frequency sensitivity of a ground-based interferometer. This noise source is a result of the thermal vibration of both the mirror suspensions and coatings.

The estimated PSD of thermal noise contributions from the suspensions and mirror coatings in the advanced LIGO detectors is plotted in figure 1.4.2. The behaviour of the PSD for the suspension has noticeable structure, with numerous peaks arising from upconversion of the resonant frequency of the suspension into higher harmonics.

1.4.1.2 Seismic noise

Seismic noise is the result of strain introduced into the interferometer through movement of the ground, which can be the result of geophysical activity, tidal activity, or anthropogenic sources of seismic noise, such as road traffic or railways. In a seismically quiet location the spectrum of

f / Hz	D / km	Sources
0.01–1.0	1000	Earthquakes, microseism
1–3	10	Anthropogenic, nearby earthquakes, wind
3–10	1	Anthropogenic, wind
10–100	0.1	Nearby Anthropogenic noise

Table 1.4.1: The principle seismic noise frequency bands, f, which affect ground-based detectors, their sources, and the distance, D, over which the band affects advanced-generation detectors.

seismic noise follows the relation [80]

$$s(f) \approx 10^{-7} f^{-2}\,\mathrm{m\,Hz^{-2}}, \tag{1.27}$$

for a frequency f.

However, the seismic environment of the detector can have a considerable effect on this noise source. Consequently, of the important considerations in choosing a site for an interferometer is the presence of seismic noise, and for this reason they are normally located far from urban areas. Table 1.4.1 summarises the approximate frequency ranges for various sources of seismic noise, and the approximate distance range over which these sources affect an interferometer. Despite this, both of the Advanced LIGO sites are affected by the presence of loud anthropogenic noise sources (LHO is affected by a nearby Department of Energy site; LLO is affected by logging activity and a nearby railway track) [101]. LLO is also strongly affected by severe storms due to its proximity to the Gulf of Mexico, especially in the microseismic band.

Seismic noise limits the sensitivity of the second generation detectors at low frequencies ($f < 50\,\mathrm{Hz}$), but it is present as a noise source across the passband of the detector. The seismic noise contains a pair of notable peaks below the 1 Hz level, one caused by ocean swell, which has a period around 4 to 30 seconds, and a second caused by standing seismic modes in the Earth

which spans the range of 30 to 1000 seconds. The presence of seismic noise below 30 Hz is still problematic for ground-based interferometers, depsite this being outside the design frequency range, due to *upconversion*, where low-frequency noise couples non-linearly into higher frequency noise.

Seismic isolation is used in detectors to reduce the noise level due to seismic activity. This takes two forms: active isolation, and passive isolation. The former is accomplished by mounting optical components on hydraulic pre-isolator systems which are controlled, via a feed-forward system, by the measurements of a seismometer. The latter is reduced by suspending the optics as a component in a pendulum system. Above the resonance of a single-stage pendulum the transfer of horizontal motion falls off as $1/f$, and vertical motion can be reduced by suspending the pendulum on a spring.

Advanced LIGO makes use of a four-stage suspension system to reduce the movement of the test mass, with the test mass forming the second stage of a two-stage pendulum which is itself suspended off two stages of cantilevered steel blades. This entire suspension system for each optic (and indeed, the entire vacuum tank containing the suspension) is placed on an isolator platform. The suspension system of Virgo follows similar principles, but involves seven stages of vertical suspension to form its super attenuators.

Seismic noise is also a source of Newtonian noise (see section 1.4.1.3) due to local mass density fluctuations as the seismic wave passes through the ground. Both the PSD of seismic and Newtonian noise are plotted in figure 1.4.3 for the advanced LIGO detectors.

1.4.1.3 Newtonian Noise

Newtonian noise, or gravitational gradient noise, is the strain produced by gravitational coupling between local mass density variations and the test masses in the interferometer. The major source of such noise comes from density fluctuations in the material surrounding the test mass, the ground below the detector. Seismic waves, especially surface waves, can produce measurable density changes which in turn affect the strength of the gravity field local to the test mass.

The spectrum of this noise is given by [102] as

$$s(f) = \begin{cases} \frac{\beta}{0.6} \frac{6\times10^{-23}}{\sqrt{\text{Hz}}} \left(\frac{10\,\text{Hz}}{f}\right)^2 & 3\,\text{Hz} \lesssim f < 10\,\text{Hz} \\ \frac{\beta}{0.6} \frac{6\times10^{-23}}{\sqrt{\text{Hz}}} \left(\frac{10\,\text{Hz}}{f}\right)^4 & 10\,\text{Hz} \lesssim f < 30\,\text{Hz} \end{cases} \tag{1.28}$$

where the β factor is site-dependent, estimated at quiet times to be 0.35 to 0.45 at LLO, and 0.35 to 0.60 at LHO.

While variations in the density of the ground are the major contribution to Newtonian noise, atmospheric and surface effects also impact the detector sensitivity. These can include the movement of clouds and aircraft in the vicinity of the detector.

1.4.2 Glitches

In addition to the sources of instrumental noise which are continuously present in interferometer data, the advanced era detectors suffer from transient non-Gaussian noise events which are known as glitch events. These can be caused by environmental phenomena, such as lightning strikes in the vicinity of the detector, or due to instrumental effects, such as fluctuations in laser power, or reflections within the beam tube. Due to their transient nature these noise events are a particular difficulty for data analysis techniques designed to identify signals from both CBC systems

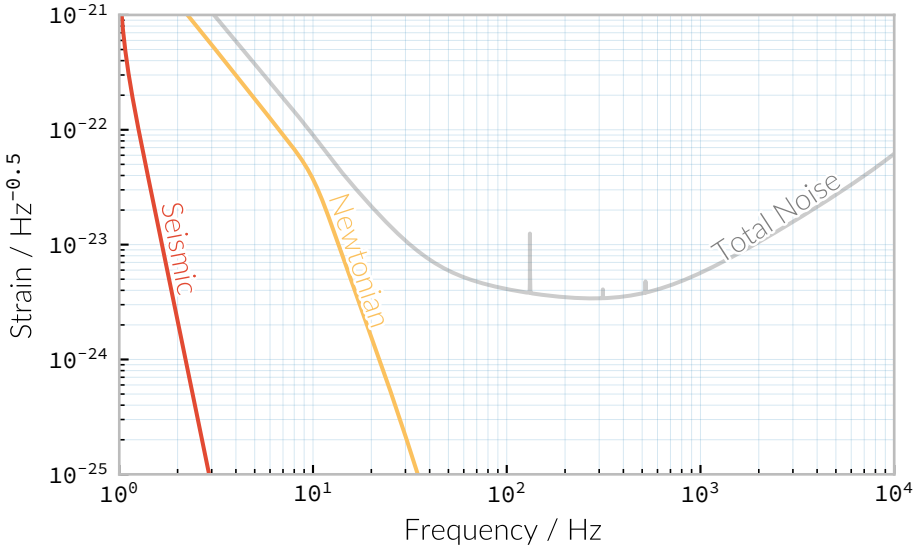


Figure 1.4.3: The contribution to the advanced LIGO PSD from seismic and Newtonian noise. These curves were calculated using the `pygwinc` library [100].

and so-called "burst" events (discussed in section 2.4). There are two major ways of addressing this problem: identifying the cause of the glitch, and making changes to the detector to reduce or eliminate their occurrence; or to produce a *veto*, a specific datum which identifies time periods where glitching is likely due to a combination of measurements from other data sources.

In order to identify the cause of any given glitch it is normally necessary to classify it; different glitch-causing phenomena will produce events with specific time-frequency morphologies. When a number of similar glitches are identified it may be possible to infer their cause with reference to the numerous sensors which monitor each detector and its site (these number on the order of 10^5 for each advanced LIGO detector). Attempts to perform this classification using a combination of human volunteers and machine learning techniques have been fruitful to date through the

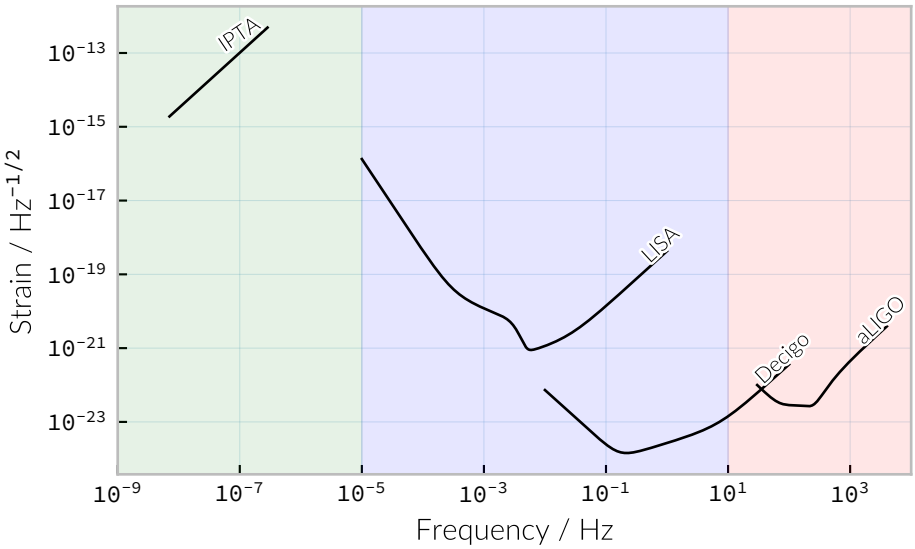


Figure 1.5.1: The gravitational wave spectrum, with a number of current and future detectors' sensitivity curves overlaid. The background colours show the regime in which each region of the spectrum can be observed, with green being the frequencies where pulsar timing is necessary, blue where space-based interferometry may be used, and pink where ground-based interferometry is currently used.

GravitySpy project [103]. Once the cause is understood either detector alteration can be planned, or a veto can be constructed with reference to data channels which *witness* the phenomena correlated with glitch production.

1.5 | A network of detectors

Generally, in order to make a confident detection of a GW the event must be observed in at least two detectors; this is principally due to the need to exclude noise sources as the source of the signal. A true GW event should be coincident (within the wave travel-time between any pair of detectors) in two or more detectors, whereas locally produced noise will appear only in the observations of a single detector, or with a time-lag which is not

physically consistent with a GW. The largely omnidirectional sensitivity of interferometric detectors further motivates the need for multiple detectors which can be used to triangulate the source of the signal in the sky.

At the time of writing the world-wide network of GW detectors was made-up of four interferometric detectors: the GEO600 detector in Germany, the advanced Virgo detector in Italy, and two advanced LIGO detectors, located in the USA states of Washington and Louisiana. The normal operation of the network omits the less sensitive GEO600 detector, and is capable of operating as a network containing all three detectors, or two detectors during periods of time where one detector is not observing.

Additional detectors are currently either being planned or are under construction which will see an increase both in the number of detectors and their geographical spread. Such an increased network should provide both an increased duty cycle (leading to a decrease in the total time when no observations are being made), and improved sky-localisation capability (improving the prospects of successful electromagnetic follow-up of GW events).

In addition to adding to the network of terrestrial detectors working in high frequencies, figure 1.5.1 demonstrates the need for detectors, such as LISA and DECIGO to be placed in space in order to observe at lower frequencies than is possible on the Earth with detectors such as advanced LIGO, in addition to the development of pulsar timing arrays to make GW observations at extremely low frequencies.

1.6 | Gravitational wave detections

Having discussed the means by which GWs may be detected, it would be remiss not to discuss the detections which occurred during the first two observing runs of the advanced detector era.

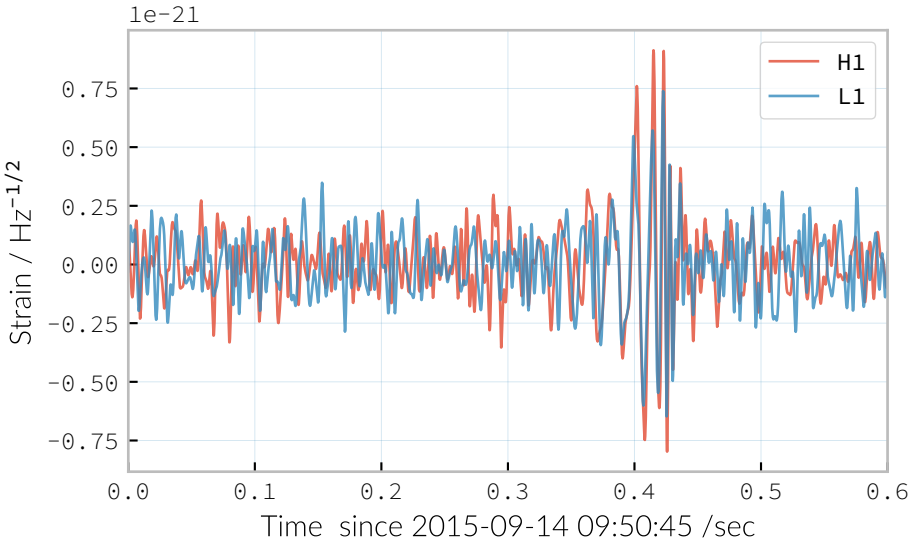


Figure 1.6.1: The data from the advanced LIGO detectors at the Livingston (L1) and Hanford (H1) observatories, which has been band-passed between 50 Hz and 250 Hz, and a comb filter has been applied to remove the 60 Hz line and its higher harmonics. The data from the Livingston detector has had a time-delay filter applied to introduce a 6.9 ms delay, representing the travel time between the detectors, and has been inverted to account for the relative orientation of the two detectors. This plot was produced using the gwpy library [104].

1.6.1 Observing run 1 and GW150914

The first detection of GWs was made on 14 September 2015 by the Advanced LIGO detectors [105] when a signal from a BBH coalescence was detected, first by the Coherent WaveBurst (cWB) burst search pipeline (which is discussed briefly in section 3.1.2), and subsequently by a number of matched-filtering pipelines designed for CBC detection. GW150914 was remarkable not only for being the first viable trigger to be detected by advanced LIGO, but also for having sufficiently high statistical significance (with a false alarm rate less than 1-in-203 000 years) that there was no reasonable doubt that it constituted a genuine GW detection; indeed, as can be seen in figure 1.6.1, the signal can be seen clearly in the whitened

data without the use of matched filtering.

The detection was made at both the LLO and LHO observatories, with a joint SNR of around 24. The event itself, a BBH coalescence between a $36^{+5}_{-4}\mathrm{M}_{\odot}$ black hole and a $29^{+4}_{-4}\mathrm{M}_{\odot}$ black hole was unexpected. Observations of black hole binaries in x-ray had not previously suggested that stellar-mass black holes this massive would exist. As a result models of stellar formation struggled to explain the evolution of black holes with these masses [106].

Two further BBH events were observed in the first observing run. GW151012, (the "second Monday event"), was initially announced as a candidate event, as it failed to exceed the 5σ significance threshold which was set for events prior to the publication of the GWTC-1 catalogue [5]. The more significant GW151226 (the "Boxing Day event") was the second confirmed detection from the advanced LIGO detectors, corresponding to a merger between much less massive black holes than GW150914 (around 14 and 8 solar masses). Unlike the first detection, GW151226 may have involved an asymmetrical system, with one black hole about twice as massive as the other. The lower masses resulted in a substantially greater amount of the inspiral waveform being in-band for the detectors, and consequently was capable of providing more stringent tests on GR than its predecessor [107].

1.6.2 Observing run 2 and GW170817

The second advanced LIGO observing run (O2) started on 30 November 2016, and finished on 25 August 2017. The advanced Virgo detector joined the run on 1 August 2017, allowing three-detector observations from kilometre-scale detectors for the first time in the advanced era. Nine detections were made during O2. These are summarised in table 1.6.1.

Of these, eight were BBH events, and one was a BNS event. The most important observation to be made during this run was of GW170817, the first detection of a binary neutron star coalescence. This event, which occurred on 17 August, was the second three-detector event (preceded only by GW170814 three days earlier), which left the community in the serendipitous situation of being able to determine the location in the sky from which the GW originated to much greater precision than previous two-detector events.

The detection of GW170817 [84] was coincident with the detection of a short gamma ray burst by the Fermi spacecraft [108]. This parallel detection of the event made GW170817 / GRB170817A the first multi-messenger GW event. Within hours of the publication of the LIGO / Virgo sky localisation an optical counterpart to the event was identified in NGC 4993 by the SWOPE Supernova Survey [109], gaining the designation AT2017gfo. The optical emission was later followed by observation of emission across the electromagnetic spectrum, including the observation of optical and ultra-violet emission (a kilonova) from the event [110].

	$E_{\rm rad}$ $/\mathrm{M}_\odot$	$L_{\rm peak}$ $/10^{56}\,\mathrm{erg/s}$	$a_{\rm final}$	$\chi_{\rm eff}$	$D_{\rm L}$ /Mpc	M_1 $/\mathrm{M}_\odot$	M_2 $/\mathrm{M}_\odot$	$\mathcal{M}$ $/\mathrm{M}_\odot$	$M_{\rm rem}$ $/\mathrm{M}_\odot$	z
GW150914	$3.1^{+0.4}_{-0.4}$	$3.6^{+0.4}_{-0.4}$	$0.69^{+0.05}_{-0.04}$	$-0.01^{+0.12}_{-0.13}$	$430.0^{+150.0}_{-170.0}$	$35.6^{+4.8}_{-3.0}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.6}_{-1.5}$	$63.1^{+3.3}_{-3.0}$	$0.09^{+0.03}_{-0.03}$
GW151012	$1.5^{+0.5}_{-0.5}$	$3.2^{+0.8}_{-1.7}$	$0.67^{+0.13}_{-0.11}$	$0.04^{+0.28}_{-0.19}$	$1060.0^{+540.0}_{-480.0}$	$23.3^{+14.0}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.0}_{-1.1}$	$35.7^{+9.9}_{-3.8}$	$0.21^{+0.09}_{-0.09}$
GW151226	$1.0^{+0.1}_{-0.2}$	$3.4^{+0.7}_{-1.7}$	$0.74^{+0.07}_{-0.05}$	$0.18^{+0.2}_{-0.12}$	$440.0^{+180.0}_{-190.0}$	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.6}$	$8.9^{+0.3}_{-0.3}$	$20.5^{+6.4}_{-1.5}$	$0.09^{+0.04}_{-0.04}$
GW170104	$2.2^{+0.5}_{-0.5}$	$3.3^{+0.6}_{-0.9}$	$0.66^{+0.08}_{-0.1}$	$-0.04^{+0.17}_{-0.2}$	$960.0^{+430.0}_{-410.0}$	$31.0^{+7.2}_{-5.6}$	$20.1^{+4.9}_{-4.5}$	$21.5^{+2.1}_{-1.7}$	$49.1^{+5.2}_{-3.9}$	$0.19^{+0.07}_{-0.08}$
GW170608	$0.9^{+0.0}_{-0.1}$	$3.5^{+0.4}_{-1.3}$	$0.69^{+0.04}_{-0.04}$	$0.03^{+0.19}_{-0.07}$	$320.0^{+120.0}_{-110.0}$	$10.9^{+5.3}_{-1.7}$	$7.6^{+1.3}_{-2.1}$	$7.9^{+0.2}_{-0.2}$	$17.8^{+3.2}_{-0.7}$	$0.07^{+0.02}_{-0.02}$
GW170729	$4.8^{+1.7}_{-1.7}$	$4.2^{+0.9}_{-1.5}$	$0.81^{+0.07}_{-0.13}$	$0.36^{+0.21}_{-0.25}$	$2750.0^{+1350.0}_{-1320.0}$	$50.6^{+16.6}_{-10.2}$	$34.3^{+9.1}_{-10.1}$	$35.7^{+6.5}_{-4.7}$	$80.3^{+14.6}_{-10.2}$	$0.48^{+0.19}_{-0.2}$
GW170809	$2.7^{+0.6}_{-0.6}$	$3.5^{+0.6}_{-0.9}$	$0.7^{+0.08}_{-0.09}$	$0.07^{+0.16}_{-0.16}$	$990.0^{+320.0}_{-380.0}$	$35.2^{+8.3}_{-6.0}$	$23.8^{+5.2}_{-5.1}$	$25.0^{+2.1}_{-1.6}$	$56.4^{+5.2}_{-3.7}$	$0.2^{+0.05}_{-0.07}$
GW170814	$2.7^{+0.4}_{-0.3}$	$3.7^{+0.4}_{-0.5}$	$0.72^{+0.07}_{-0.05}$	$0.07^{+0.12}_{-0.11}$	$580.0^{+160.0}_{-210.0}$	$30.7^{+5.7}_{-3.0}$	$25.3^{+2.9}_{-4.1}$	$24.2^{+1.4}_{-1.1}$	$53.4^{+3.2}_{-2.4}$	$0.12^{+0.03}_{-0.04}$
GW170817	> 0.04	> 0.1	< 0.89	$0.0^{+0.02}_{-0.01}$	$40.0^{+10.0}_{-10.0}$	$1.46^{+0.12}_{-0.1}$	$1.27^{+0.09}_{-0.09}$	$1.186^{+0.001}_{-0.001}$	< 2.8	$0.01^{+0.0}_{-0.0}$
GW170818	$2.7^{+0.5}_{-0.5}$	$3.4^{+0.5}_{-0.7}$	$0.67^{+0.07}_{-0.08}$	$-0.09^{+0.18}_{-0.21}$	$1020.0^{+430.0}_{-360.0}$	$35.5^{+7.5}_{-4.7}$	$26.8^{+4.3}_{-5.2}$	$26.7^{+2.1}_{-1.7}$	$59.8^{+4.8}_{-3.8}$	$0.2^{+0.07}_{-0.07}$
GW170823	$3.3^{+0.9}_{-0.8}$	$3.6^{+0.6}_{-0.9}$	$0.71^{+0.08}_{-0.1}$	$0.08^{+0.2}_{-0.22}$	$1850.0^{+840.0}_{-840.0}$	$39.6^{+10.0}_{-6.6}$	$29.4^{+6.3}_{-7.1}$	$29.3^{+4.2}_{-3.2}$	$65.6^{+9.4}_{-6.6}$	$0.34^{+0.13}_{-0.14}$

Table 1.6.1: The events from the first two advanced-era observing runs. The data in this table is derived from the first gravitational wave transient catalogue, GWTC-1 [5]. $E_{\rm rad}$ is the total GW energy radiated as a result of the event; $L_{\rm peak}$ is the event's peak GW luminosity; $a_{\rm final}$ is the total spin of the remnant black hole; $\chi_{\rm eff}$ is the effective spin of the CBC system; $D_{\rm L}$ is the luminosity distance to the source; M_1 and M_2 are the masses of the two compact objects; $\mathcal{M}$ is the chirp mass of the system; $M_{\rm rem}$ is the mass of the remnant, and z is the redshift of the source.

Epoch	LIGO (Mpc)	Virgo (Mpc)	KAGRA (Mpc)
Early	40 - 80	20 - 65	8 - 25
Mid	80 - 120	68 - 85	25 - 40
Late	120 - 170	85 - 155	40 - 140
Design	190	125	140

Table 1.7.1: The anticipated sensitivities of the various second-generation detectors throughout their development, measured in terms of the BNS horizon distance, which represents the average maximal distance at which the signal from a binary neutron star coalesence could be observed. This table was adapted from the information in [111].

1.7 | Future observing scenarios

The work in this thesis will consider the state of GW detection in the observational era, starting in the early observational period: the first two observing runs of the advanced LIGO detectors, and the first observing run of the advanced Virgo detector; looking ahead to future observing runs involving a larger network of GW detectors, including Kagra and an additional advanced LIGO detector located in India.

The development of the advanced detectors is still on-going; sensitivity improvements are normally made incrementally during periods when the detectors are taken offline for extended periods of time. This phased approach means that the sensitivity of the detectors, and consequently the detector network, will improve in subsequent observing runs. In table 1.7 these are summarised; the *early* scenario equates approximately to the O1 run for advanced LIGO, and the O2 run for advanced Virgo. Similarly, the *mid* and *late* scenarios correspond approximately to O2 and O3 for advanced LIGO.

The first two observing runs have provided some information about the rate of the events which produce detectable GWs, allowing better constraints to be placed on anticipated observed event rates as the detectors

continue to develop over the next decade.

2 Astrophysical sources of gravitational waves and their waveforms

This chapter discusses the astrophysical sources of gravitational waves, and the form that the gravitational waves take, which we expect to detect as signals. Section 2.1 introduces continuous sources of gravitational waves, such as gravitational wave pulsars. Section 2.2 provides a brief overview of the stochastic gravitational wave background. Section 2.3 discusses compact binary coalescence events, which include binary black hole and binary neutron star merger events. This section contains discussion of the waveform of these events in section 2.3.2, and the numerical relativity techniques used to produce the most accurate waveforms available in section 2.3.2; a discussion of the available catalogues of these data is contained within section 2.3.4. Because of the computational expense of producing these waveforms a number of analytical approximant waveforms exist, and I discuss some of these in sections 2.3.5 and 2.3.6.

There are many conceivable astrophysical situations where computing the waveform beforehand is likely to be impractical, or impossible. As a result a number of modern searches attempt to search for *unmodelled* or poorly-modelled signals. Transient signals of this type are normally called

burst signals. Section 2.4 provides an overview of burst-like waveforms, and some of the potential sources of these signals.

Finally, in section 2.4.8 I provide discussion and results from a study I have conducted into the detectability of burst signals resulting from encounters between black holes which either result in glancing, parabolic encounters, or radiation driven capture and a subsequent merger event.

The morphologies of gravitational wave (GW) signals can be divided roughly into three categories [52], which correspond approximately to the classifications of the astrophysical sources which produce them.

Continuous signals are expected to be produced by sources over long periods of time. The primary source of continuous sources for Laser interferometer gravitational-wave observatory (LIGO) are expected to be GW pulsars, but in detectors which are sensitive at lower frequencies, such as the proposed Laser Interferometer Space Antenna (LISA) mission, the radiation from inspiralling binary systems should also be detectable.

Stochastic signals are expected to constitute a background of GWs, produced by the black holes at the centres of galaxies [112, 113, 114, 115], and from the Universe's inflationary period [116].

Transient signals are strong *bursts* of GWs over a period of seconds or less. The sources of transient signals are normally further sub-divided. *Unmodelled* sources, where there are insufficient theoretical models to use matched filtering techniques to search for signals in detector data. *Modelled* sources, which currently encompass compact binary coalescence (CBC) sources are sufficiently well understood, in contrast, to allow the use of matched filter searches. These are sources which are primarily expected in the advanced LIGO passband, with compact binary coalescences and supernovae being major targets for burst searches in the advanced observing runs, however there are prospects for burst sources in the LISA regime, for example from hyperbolic encounters between compact objects and stars or other compact objects [117, 118, 119, 120, 121, 122, 123, 124].

2.1 | Continuous wave sources

The discovery of pulsars by Bell and Hewish in 1967 was an unexpected discovery for radio astronomy—objects which produce beams of radiation, and rotate rapidly. Not long after their discovery it became apparent that they were a specific form of neutron stars (NSs) — the tightly-packed remnant of a massive star which has ended its life as a supernova.

Any rotating mass quadrupole (see equation 1.9) will produce GWs as it rotates (see equation 1.8), and so it follows that a dense, massive object, such as a neutron star, will produce GWs as they rotate, if they possess any mass asymmetry. Further, thanks to their highly stable rotation speed, the GW emission from NSs ought to be produced continuously at a very stable frequency. As GW emission occurs in the quadrupole [(2,2)] mode [1], this emission should be at twice the rotational frequency of the NS.

The rotation frequency of most pulsars is well-measured by radio observations, making these attractive prospective sources for ground-based GW detectors, such as LIGO and Virgo. Indeed, targeted searches for GWs from pulsars at twice their rotation frequency have been conducted since the "initial" detector period, using data from LIGO, GEO600, and later Virgo (which are summarised in [48]), and continued into the advanced era [125, 126].

Pulsar spins are known to decelerate over time, through a process known as spin-down. This process is often attributed to energy loss through gravitational radiation. The spin-down limit of a pulsar is the GW strain which corresponds to this scenario, where the entirety of the energy being lost is radiated as GW, and is defined as

$$h = \left(\frac{5}{2}\frac{GI_{zz}|\dot{f}|}{c^3 d^2 f}\right)^{\frac{1}{2}}, \tag{2.1}$$

[1] In general relativity (GR), at least.

for a pulsar with a spin frequency f, moment-of-inertia I_{zz}, at a distance d from the observer.

To date no continuous GW source has been detected, but the failure to measure GWs from known pulsars has allowed tight limits to be placed on the ellipticity of 222 known radio pulsars [127], with the tightest limit placed on PSR J0711−6830: $h = 1.2\times10^{-8}$.

Searches made for pulsars at around twice their rotation frequency were augmented with targeted searches for emission at the rotation frequency following the second observing run [127]. This is possible if the rotating neutron star is either biaxial or triaxial, and exhibits free precession. Emission at the rotation frequency may also be possible in NSs with pinned superfluid interiors [128].

Continuous wave searches can also be used to test GR, and to place limits on the parameters of alternative theories of gravity. Searches for non-tensorial polarisations of GW were conducted on data from the first advanced observing run [129] for all six potential polarisations allowed in general metric theories.

In comparison to transient signals, the waveforms for continuous wave sources are generally (semi-)analytical; the waveform model used for the search for the (2,2)-mode emission in [127] for example has the form

$$\begin{aligned} h_{22}(t) = -C_{22}\big[&F_{+}^{D}(\alpha,\delta,\psi,t)(1+\cos^{2}\imath)\cos(2\Phi(t)+\Phi_{22}^{C}) \\ &+2F_{\times}^{D}(\alpha,\delta,\psi,t)\cos\imath\sin(2\Phi(t)+\Phi_{22}^{C})\big], \end{aligned} \tag{2.2}$$

for C_{22} the amplitude of the wave, and Φ_{22}^{C} its initial phase at some specific time. $\Phi(t)$ is the rotational phase of the source, and $\imath$ is the inclination of the source to the observer. In contrast to the signal from transient sources, continuous waves can be observed over long periods of time, and the observed signal will be convolved with the antenna pattern of the

detector, $F^D_{+,\times}$, which varies with the sky position of the source, (α, δ), the polarisation angle of the source, and thanks to the relative movement of the source and the detector, time, t.

2.2 | Stochastic backgrounds

In addition to transient and continuous sources of GWs, which originate from specific locations in the sky, we expect that a *background* of GWs should be observable throughout the sky (with an approximately isotropic distribution). The GW background is expected to cover the entire frequency range at some level, from extremely low frequencies (around an inverse Hubble-time) to frequencies exceeding 10^{14} Hz.

The GW background is approximately analogous to the cosmic microwave background (CMB). This is the near-isotropic electromagnetic emission discovered in 1964 as constant background radio emission across the sky [130], which originates from the epoch of recombination, when atoms started to form, and the universe became optically thin.

This GW would be the result of numerous unresolved and weak signals, and as a result this background would be continuous and approximately isotropic [131], it is likely to fall into the part of the measured GW data which is treated as noise by the majority of analyses. The noise produced by the detector will be greater than this signal, and so detection of a background is also reliant on correlations between a network of detectors. This reliance on correlations between a network of detectors significantly affects the sensitivity of the detector network to background sources [132].

The level of anisotropy in the CMB implies that the universe today must be very-nearly flat, and since any curvature would increase as the universe undergoes metric expansion, this would suggest that the early universe was even flatter. This poses a dilemma, as a flat universe requires

the energy density of the universe to be equal to a critical energy density, with a small deviation becoming exagerated over time. Inflation is an attempt to address this problem, by suggesting that the universe expanded extremely rapidly early in its evolution, through the effect of some scalar field. Such a scalar field would be subject to quantum fluctuations, and tensor fluctuations would be expected to produce GWs [116]. GW production is not predicted in the early universe by non-inflationary models, and so discovery of an inflationary GW background would be strong evidence for the inflation model [133].

First-order phase transitions, which occur when the thermodynamic properties of a system are discontinuous (such as the sudden, discontinuous change in the entropy and volume of a liquid as it boils) could also be responsible for the production of GWs in the early universe. A number of phase transitions are believed to have occured as, for example, the strong and electroweak forces decoupled [134].

Cosmic strings may also be a viable source of background radiation [135]. These are topolgical defects which are caused by symmetry phase transitions in a number of grand unified theories. As these defects move they interact with each other to form kinks and cusps, which can be the source of bursts of gravitational radiation. Over a sufficiently long period the signals from these events can superimpose to form apart of the GW background.

The inspiral of the very large number of compact binary systems in the universe will also superimpose to contribute to this background radiation. These systems include galactic white dwarf binaries, which are expected to produce such a strong signal that they will limit the sensitivity of the LISA detector, as well as binary black hole (BBH) and binary neutron star (BNS) systems [136, 137].

Searches for a stochastic background have been made using the data

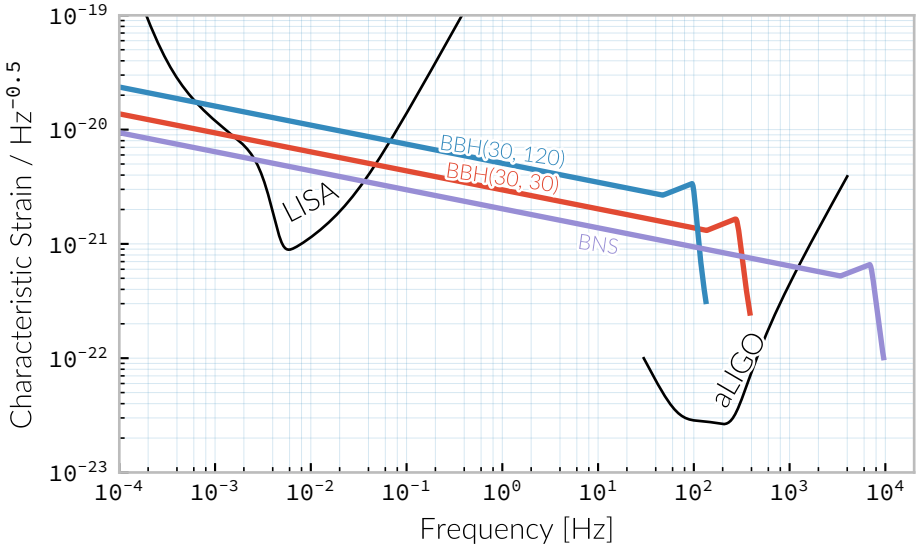


Figure 2.3.1: The frequency spectrum of two types of compact binary coalescence—a binary neutron star coalescence, and a binary black hole coalescence—alongside the design sensitivity power spectrum of the Advanced LIGO detector at its design sensitivity.

from the advanced LIGO detectors during their first observing run [138]. To date no evidence of a background have been identified, which has allowed an upper limit to be placed on its strength. Additional limits have been placed thanks to astrometric measurements of active galactic nuclei using radio data and the first GAIA data release [139], and through pulsar timing arrays [140, 141].

2.3 | Compact Binary Coalescences

The moment of inertia tensor of a two-body system will lead to non-spherical motion within the system, which will in turn produce gravitational radiation, and gradual orbital decay [142]. This effect was first observed in the Hulse-Taylor pulsar [6, 7], a system containing two neutron stars—one of which is a pulsar—which, through precise pulsar timing

measurements, were inferred to be inspiralling, and producing GWs. This observation was the first strong, indirect evidence for the reality of GW production.

The binary system will continue to lose energy via gravitational radiation until it reaches its *innermost stable circular orbit*, after which the objects will merge, and eventually coalesce. These coalescence events are powerful sources of GWs, and the chirp produced by the coalescence may be sufficiently luminous to be detected by current, advanced detectors. The frequency spectra of both a binary neutron star and a binary black hole coalescence are plotted in figure 2.3.1, compared to the sensitivity curve of the advanced LIGO detectors. Attempts to detect signals from such coalescing systems were also made during the initial run of the detectors, prior to their being upgraded to advanced LIGO and advanced Virgo [50]. These would be characterised by the distinct pattern of the pseudo-sinusoidal inspiral waveform, followed by a bright burst of radiation, and then a sinusoidal *ringdown* as the post-coalescence remnant vibrates [52]. Binary coalescences are thus classified as transient, or burst sources.

The potential objects which may be involved in an binary coalescence observable by the current generation of ground-based detectors are black holes and neutron stars: both compact objects. In the future white dwarf binaries may also be observable at lower GW frequencies, and these systems are expected to be much more abundant than either neutron star or black hole binaries, but their emission lies within the passband of LISA —a planned space-based GW observatory. These never reach a last stable orbit, as it lies within their physical diameter, and so the inspiral component of the waveform is the principle source of GWs. These are expected to be so numerous in LISA results [60] that entirely new statistical methods will be needed to process the observations, and to allow observations of other phenomena to be made in their background.

Supermassive binary black holes (SMBBHs) are believed to collide and merge as part of the merging processes of galaxies. Again these binaries should produce signals within the passband of LISA [143], and should be so spectacularly strong that they are visible in the LISA data without the *matched filtering* techniques which are required to extract other signals from the data [52]. The observation of these objects would provide much-needed information about the evolution of galaxies and of super-massive black holes.

Inspiralling compact binaries can act as a cosmological distance measure: they have two parameters, their period, and the rate at which that period changes (which is calculated by measuring the chirp mass of the system) which characterise the system, and the amplitude of the GWs produced is dependent only on the chirp mass of the source, and the distance from the observer to the object. As a result it is possible to determine the distance to an inspiralling system simply by determining the chirp mass and measuring the brightness of the event [144]. This would provide an additional means of measuring cosmic acceleration, and, in the LISA era, this would allow the measurement of acceleration at high redshift using high-mass binary black holes.

2.3.1 Dynamics of compact binaries

The dynamics of binary systems are well-understood in Newtonian mechanics, where the two-body problem can be reduced to a pair of independent one-body problems. In contrast no exact solutions have been found to this problem in GR; while the Schwarzschild solution [145] is sufficient for some situations where the mass of one of the two bodies is much smaller than the other (where the problem is effectively a one body problem) it is insufficient for systems such as BBHs.

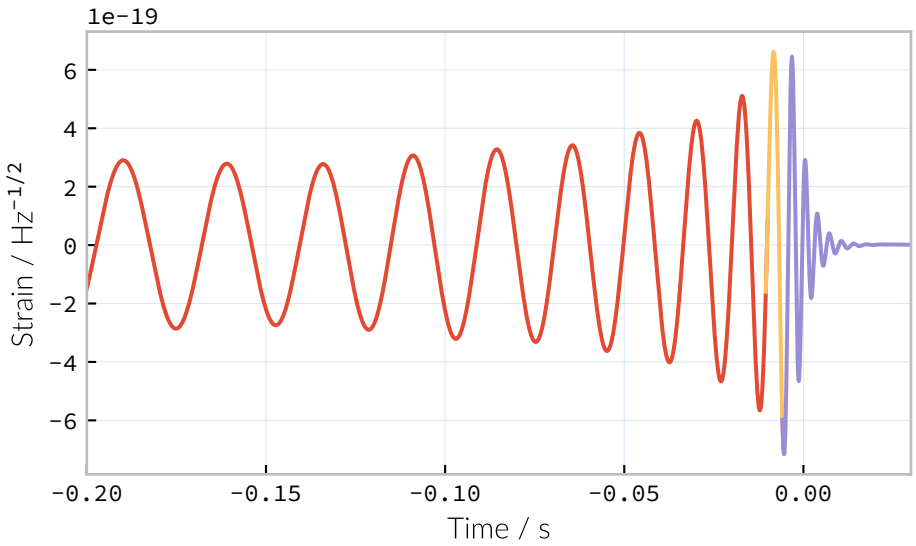


Figure 2.3.2: The waveform of a BBH from the inspiral (highlighted in red), to the merger (yellow), and the ringdown (purple).

When the two component bodies of the system are at large separation (and their local velocities are much smaller than the speed of light) a post-Newtonian (PN) expansion can be used. In this regime the two objects are treated as point-particles with slow internal dynamics. The PN correction to the Newtonian limit on the order $\mathcal{O}(1/c^n)$ is generally called the $(n/2)$-PN order.

2.3.2 The compact binary waveform

The gravitational waveform for a compact binary system can be split into three broad periods, which are each associated with the dominant energy loss mechanisms within the system. These are illustrated on a plot of the time-domain waveform of a BBH signal in figure 2.3.2. The first, and longest stage of the binary's evolution is the *inspiral*. Gravitational radiation carries energy out of the binary system, causing the orbit to

slowly decay.[2] For the majority of the inspiral the GWs produced have a very low amplitude, and are to weak to be detected by the current generation of detectors, however this amplitude increases as the radius of the orbits decrease. Eventually this amplitude becomes observable, for a period ranging from minutes (in the case of BNS events [84]), to fractions of a second (for most BBH events [105]).

As the binary reaches its innermost stable circular orbit the system evolves from the inspiral period to the merger. At this point the two black holes *plunge* towards each other, and then coalesce. This period contains the peak emission of GWs.

Finally, the single black hole which remains will radiate energy through the *ringdown* period, during which the black hole oscillates, radiating energy until it becomes a stable Kerr black hole.

2.3.3 Numerical relativity

The study of compact binary systems using GWs relies on solving the relativistic two-body problem; the classical, Newtonian solutions to this problem are Keplerian orbits, however post-Newtonian gravity requires that a mass with orbital angular momentum loses energy in the form of GWs.

The field of numerical relativity (NR), while now capable of producing accurate waveforms for a wide variety of initial BBH conditions, had a lengthy period of development. As recently as 1999 Brügmann [146] notes that "the binary black hole problem is essentially unsolved". The major stumbling-blocks for NR were specific to GR. The first of these is the gauge freedom of the theory, which generally makes specifying

[2] This in fact occurs in *all* orbits, however most objects will not get close enough that the current generation of detectors will be able to observe the low-amplitude radiation produced by such systems. In the future, however, inspirals of objects such as white dwarf binaries are expected to be noise sources for space-based detectors, such as LISA.

a numerical coordinate grid on the simulated spacetime impossible in advance. As a result an effective method of producing such a coordinate grid during the evolution of the numerical simulation must be employed, which avoids the introduction of coordinate singularities. Coordinate singularities represented the second major challenge to the field; in the case of BBH spacetimes these are typical features. Additionally, finding a formalism for the representation of the field equations throughout the evolution of the simulation which would remain stable had proved challenging.

By the mid-2000s a number of breakthroughs occured. In 2004 Pretorius [147] introduced the "generalised harmonic coordinate" formalism which remained stable into the evolution of the BBH merger. This was followed [148] by the demonstration of an NR simulation which evolved the BBH through the inspiral and merger to the ringdown produced in this formalism. Late in 2005 Campanelli *et al.* [149] demonstrated the use of an algorithm which overcame the difficulties of coordinate singularities inherent in black hole simulations. These had previously been overcome through the "excission" of the black hole, where a boundary was placed inside the black hole event horizon, excising its interior, containing the singularity, from the computational domain. Instead, their technique employed "punctures", where the poles which represented the black holes were factored-out analytically, allowing the production of accurate and complete waveforms. Also in 2005 Baker *et al.* [150] developed a technique for extracting the BBH waveform directly from the outer region of the simulation, based on the work of Fiske *et al.* [151].

The covariant nature of the Einstein field equationss (EFEs) makes choosing a frame of reference in which to evaluate the metric difficult, and complicates the process of defining an initial value problem to solve. To get around this the EFEs are often decomposed into a $(3+1)$-dimensional foliation, in which the dynamics at each time slice can be solved. A large

range of scales must be resolved within NR simulations in order to model both the behaviour of spacetime close to the merging system and at the location that the GW is extracted. This range of scales makes evolving an NR simulation computationally burdensome, even with techniques such as adaptive mesh refinement which aim to make this process efficient. A recent review by Lehner and Pretorius [152] of the techniques involved in running NR simulations summarises a number of the techniques which are used to make NR tractable.

The complexity of NR simulations has lead to their adoption of parallelisation technology for multiprocessing and message-passing between processes, however these simulations can still require around a month to produce on computing clusters containing thousands of processors.

A number of codes are used to produce NR waveforms for the advanced-era GW detectors:

SPEC The Spectral Einstein Code (SpEC) is a NR method which leverages spectral methods during the evolution of the black hole spacetime in BBH simulations, in an attempt to circumvent instabilities which are present when using finite difference methods [153].

The code is capable of generating the merger and ringdown component of the GW waveform for a generic BBH configuration [154].

BAM The bi-functional adaptive mesh (BAM) code uses a modified Baumgarte-Shapiro-Shibata-Nakamura (BSSN) regime [155, 156, 155, 146].

MAYA The MAYA code is based on the BSSN formalism with a moving puncture gauge condition [157].

2.3.4 Catalogues

A number of catalogues of NR computed BBH waveforms are readily available; two of the largest originate from the relativity group at Georgia Institute for Technology and the SXS collaboration.

2.3.4.1 Georgia Tech Waveform catalogue

The Georgia Tech waveform catalogue [157] is composed of 452 waveforms which were generated using the MAYA NR code at the Centre for Relativistic Astrophysics at Georgia Institute of Technology. The catalogue includes both non-spinning simulations for quasi-circular systems with mass-ratios $q \leq 15$, and precessing quasi-circular systems with $q \leq 8$. Within the set of waveforms derived from spinning systems are two subsets: aligned-spin, where the spin axis of each black hole is parallel to the orbital angular momentum, $\boldsymbol{L}$; and precessing, where the spin axes are not parallel to $\boldsymbol{L}$. The distribution of BBH parameters for the waveforms in the catalogue are plotted in the corner plot of figure 2.3.3.

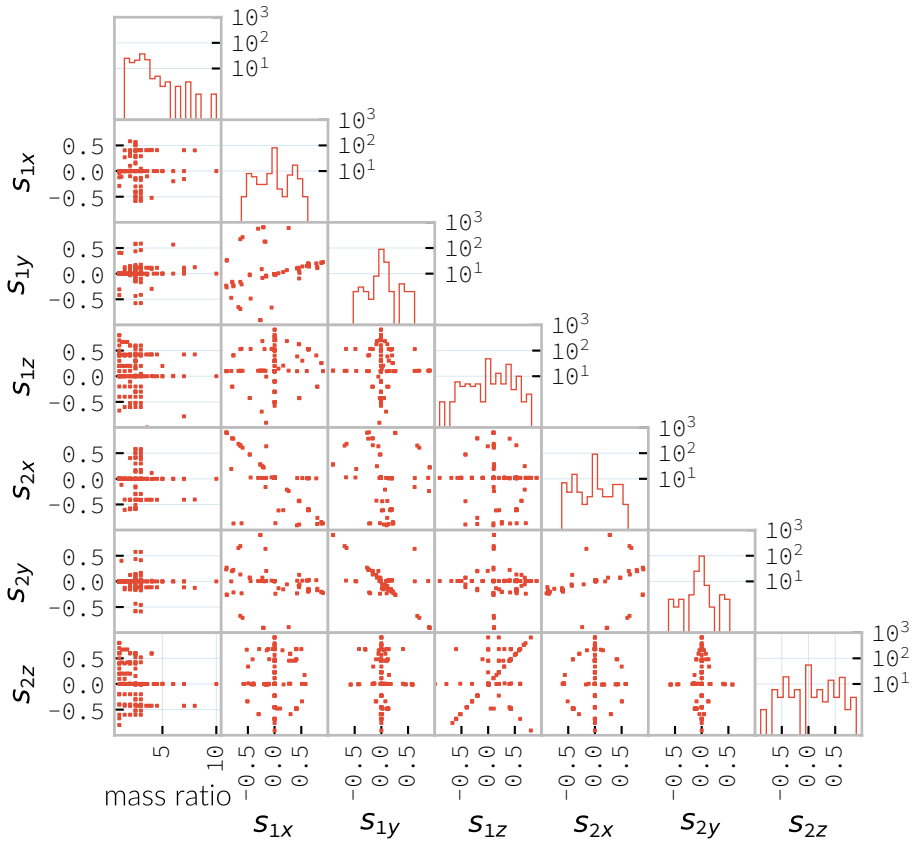
2.3.4.2 SXS waveform catalogue

The SXS waveform catalogue [158, 159, 160, 161, 162, 163] includes over 400 waveforms, both spinning and non-spinning, generated using SpEC. The coverage of this catalogue is shown in the corner plot of figure 2.3.4.

2.3.5 Analytical approximants

The impossibility of producing enough NR waveforms to densely cover even the two dimensional parameter space of non-spinning BBH systems has lead to the development of algorithms capable of producing approximations of the waveform across the parameter space. While PN approximants provide a powerful approximation to the waveform in the inspiral phase,

Figure 2.3.3: The coverage of the Georgia Tech catalogue over the intrinsic physical parameter space of BBH systems.

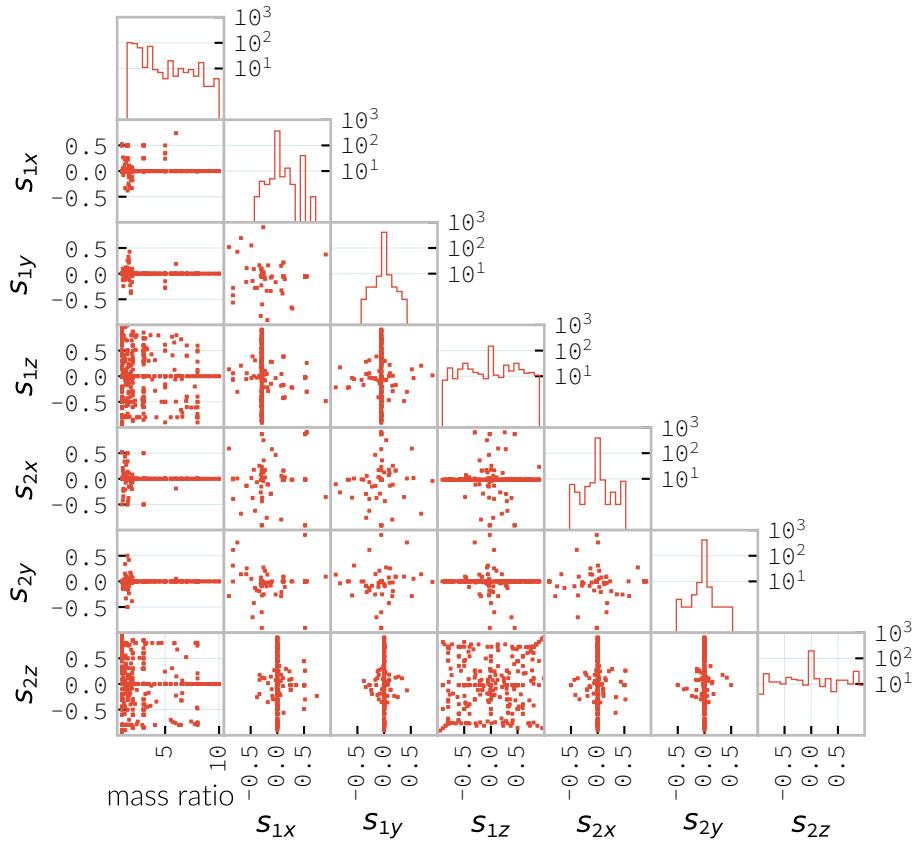


as the characteristic velocity of the binary approaches the speed of light the PN expansion will lose accuracy, and an alternative method for approximating the waveform around the merger is required. There are currently two major implementations of such approximants; the IMRPhenom family, and the SEOBNR family of approximants.

2.3.5.1 IMRPhenom

The IMRPhenom models [164] take advantage of the three-component structure of BBH signals (see 2.3.2); calibration waveforms for the models are produced by a NR simulation. The calibration waveforms which are

Figure 2.3.4: The coverage of the SXS waveform catalogue over the intrinsic physical parameter space of BBH systems.



produced by NR are short, and the inspiral is normally calculated for only the last few cycles of the binary. In order to make a longer waveform these NR waveforms are joined to a PN inspiral waveform (since the PN is known to be a good approximation for this part of the waveform) in a process called "hybridisation". For hybridisation to be effective the PN and NR waveforms must be well-matched. This match is determined by their integrated squared absolute difference,

$$\delta = \int_{t_1}^{t_2} \left| \mathsf{h}^{\mathrm{PN}}(t, \boldsymbol{\mu}) - a\mathsf{h}^{\mathrm{NR}}(t, \boldsymbol{\mu}) \right|^2 \mathrm{d}t, \tag{2.3}$$

with h^{NR} an NR waveform, h^{PN} a PN waveform evaluated at the same parameters, a is an amplitude scaling factor, and $\boldsymbol{\mu}$ a vector of extrinsic

parameters, $\boldsymbol{\mu} = \{\phi_0, t_0\}$, the initial phase and start time of the waveform, respectively [165].

The resulting hybridised waveforms are then parameterised in the Fourier domain. These *phenomenological* waveforms, $u(f)$ take the form

$$u(f) = A(f)\exp(i\Psi(f)), \tag{2.4}$$

for Ψ the phase, and with a piecewise function describing the amplitude, A as a function of frequency, f:

$$A(f) = C\begin{cases} (f/f_{\text{merge}})^{-7/6} & \text{if } f < f_{\text{merge}} \\ (f/f_{\text{merge}})^{-2/3} & \text{if } f_{\text{merge}} < f < f_{\text{ring}} \\ w\mathcal{L}(f, f_{\text{ring}}, \sigma) & \text{if } f_{\text{ring}} < f < f_{\text{cut}}, \end{cases} \tag{2.5}$$

where f_{merge}, f_{ring}, and f_{cut} are respectively the initial merger frequency, initial ringdown frequency, and the cutoff frequency of the template. $\mathcal{L}$ is a Lorentzian distribution of width σ, and w is a normalisation constant which describe the quasi-normal mode frequencies, and C is a numerical constant (details of these parameters can be found in [165]).

The effective phase, Ψ, expanded in powers of f, is

$$\Psi = 2\pi f t_0 + \phi_0 + \sum_{k=0}^{7} \phi_k f^{(k-5)/3}, \tag{2.6}$$

with ϕ_0 the phase offset, each of the ϕ_k values phase parameters, t_0 the arrival time of the waveform.

The amplitude and phase parameters of these phenomenological waveforms are then determined by fitting the model to around thirty hybridised waveforms. Finally, the best-matching amplitudes and phases for the phenomenological waveforms are fitted to the physical parameters of the binary in order to produce a physically parameterised model.

The first model to take this approach, IMRPhenomA, was calibrated only against non-spinning hybrid waveforms. Further development pro-

duced the IMRPhenomD model [166], which is calibrated against 19 hybrid PN - NR waveforms (a mixture of public SXS and BAM-derived waveforms) to produce aligned-spin spinning waveforms. The IMRPhenomD model is then verified against 29 additional hybrid waveforms.

The IMRPhenomP series of waveform models (the most recent of which is version 3 [167]) add the ability to model precession effects within the waveform; for versions 1 and 2 this was limited to single-spin effects, but version 3 has been designed to allow for generic BBH systems. In order to introduce the effects of precession into the waveform, `IMRPhenomPv1` and `IMRPhenomPv2` built on the non-precessing waveforms from the IMRPhenomC and IMRPhenomD families, respectively, and then added the modulations produced in the waveform by orbital precession. For these first two versions the precession angles were calculated by a frequency-domain expression which assumed a single-spin system, under the stationary phase approximation [3], which is not strictly valid outwith the inspiral phase. `IMRPhenomPv3` uses a two-spin model developed by Chatziioannou *et al.* [169] in order to allow for the calculation of precession angles in generic BBH systems.

2.3.5.2 Effective one-body

An alternative approach to the phenomenological fitting of the IMRPhenom algorithms is the effective one body (EOB) approach. The EOB approach [170, 171, 172] maps the dynamics of two compact objects into that of a single test particle moving in a deformed Kerr metric. In contrast to the piecewise approach to building the waveform taken in the IMRPhenom model (see section 2.3.5.1), the EOB approach constructs the entire waveform in a single process [173]. The waveform is constructed by

[3] The stationary phase approximation is found to provide sufficient accuracy for the matched-filtering processes which are common in GW data analysis. [168]

assuming that the merger is short but with a broad range of frequencies; this section of the waveform is built by attaching the signal from a plunge signal to quasinormal modes.

Similarly to IMRPhenom, the EOB derived waveforms are calibrated against a number of NR derived waveforms. For the non-spinning model, `EOBv2` this involved five waveforms produced by the `SPEC` code.

2.3.6 Numerical relativity surrogate models

Recently, an entirely different approach to approximating the BBH waveform has started to emerge, based on *surrogate modelling*. These models attempt to directly model NR waveforms without introducing phenomenological assumptions, or approximations to GR, and take what might be considered a *data-driven*, or statistical approach to the problem. While the ability to abandon these assumptions and approximations is attractive, it comes at the expense of requiring a large number of NR waveforms with which to condition the model. To date, there have been two approaches to building such models: those using spline regression, and those using *Gaussian process regression*. This section will contain a broad overview of the former, but a thorough discussion of the latter will be given later in this work (in chapter 6).

The NRSur family of surrogate models, developed by Blackman *et al.* [174, 175, 176] employ spline interpolation to waveforms generated by the `SpEC` NR code. The two analysis-ready versions of this model, NRSur4d2s and NRSur7d2s are capable of producing waveforms for systems with a mass-ratio < 2 and an effective spin-parameter < 0.8. In contrast to phenomenological models, the NRSur models are currently capable of producing only a small number of cycles of the waveform, being limited by the length of the NR waveforms off which they are conditioned. Recent efforts

have been made, however, to produce similar surrogate models which are conditioned on hybridised waveforms [159]. The number of waveforms required to produce the surrogate model is also considerably larger than those required for the phenomenological models, with NRSur7d2s being conditioned on 744 NR waveforms.

An alternative approach to spline surrogate models, which rely on Gaussian process regression (see chapter 6) has recently been shown to be viable [2]). The development of these models is discussed in detail in chapter 7.

2.4 | Unmodelled and poorly modelled transient sources

While CBC searches which are designed to identify well-known signal morphologies in detector data, *burst* searches are intended to identify signals either where there is no prior knowledge of the signal morphology, or where that morphology is poorly modelled.

2.4.1 Parameterisation of burst signals

While signals from well-defined astrophysical systems, such as CBC signals, can be parameterised according to the intrinsic and extrinsic properties of the generating system, burst signals do not have a well-defined physical model. As a result we must define a number of parameters based purely on the properties of the signal.

The first of these is the *Characteristic squared amplitude*, $||h^2||$. This quantity is frequently referred to as the root square sum strain (hrss).

Definition 2.4.1. The characteristic squared amplitude, $||h^2||$, is defined

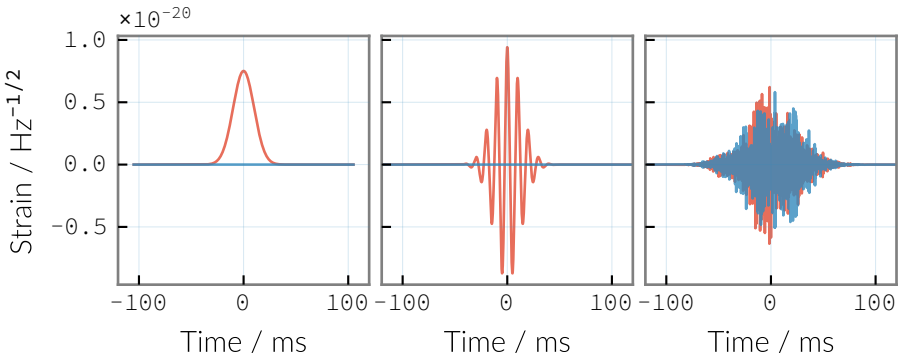


Figure 2.4.1: The three unmodelled burst "waveforms" which are typically considered by burst analyses; Gaussian-like bursts, Sine-Gaussian bursts, and White noise bursts each depicted in the time domain, with both the plus polarisation (red) and cross polarisation (blue) depicted.

as

$$||h^2|| = \int_{\infty}^{\infty} |h(t)|^2 \, \mathrm{d}t = \int_{\infty}^{\infty} |\tilde{h}(f)|^2 \, \mathrm{d}f, \tag{2.7}$$

for $h(t)$ and $\tilde{h}(f)$ respectively the strain in the time, t, and frequency f representations [177].

For bursts which are well-localised in time we can also define a central time and a duration.

Definition 2.4.2. The central time, t_0, is defined as

$$t_0 = \int_{-\infty}^{\infty} t \frac{|h(t)|^2}{||h^2||} \, \mathrm{d}t, \tag{2.8}$$

and the duration, σ^2 is defined

$$\sigma^2 = \int_{-\infty}^{\infty} (t - t_0)^2 \frac{|h(t)|^2}{||h^2||} \, \mathrm{d}t. \tag{2.9}$$

Equivalently, for bursts well-localised in frequency we can define a central frequency and a bandwidth.

Definition 2.4.3. The central frequency, f_0, is defined as

$$f_0 = \int_{-\infty}^{\infty} f \frac{|h(f)|^2}{||h^2||} \, \mathrm{d}f, \tag{2.10}$$

and the duration, b^2 is defined

$$b^2 = \int_{-\infty}^{\infty} (f - f_0)^2 \frac{|h(f)|^2}{||h^2||} \, \mathrm{d}f. \tag{2.11}$$

We can also define a *quality factor*, Q for a burst.

Definition 2.4.4. The quality factor, Q, of a burst signal is defined as

$$Q = f_0^2/b^2. \tag{2.12}$$

2.4.2 Gaussian bursts

Perhaps the simplest conceivable model of a burst of GWs is one where energy is emitted across a broadband range of frequencies over a fixed period of time, with a smooth rise and decay in amplitude. Such a source can be modelled as with a Gaussian function, and may be a suitable model for broadband sources, such as the core-bounce during a core-collapse supernova (SN).

In searches the model for such a signal is

$$h(t) = A \exp\left(-\frac{(t - t_0)^2}{2\sigma^2}\right), \tag{2.13}$$

for a strain h at time t, with an amplitude A, central time t_0 and duration σ.

An example of a Gaussian burst waveform ($\sigma = 0.01\,\mathrm{s}$, $A = 1 \times 10^{-21}$, and $t_0 = 100\,\mathrm{s}$) is plotted in the left column of figure 2.4.1. In this figure the two polarisations of the signal are plotted, with only the plus polarisation containing GW power for this morphology.

2.4.3 Sine-Gaussian bursts

In addition to searching for broadband, time-constrained bursts of GW energy, some sources are expected to produce GWs which are in a confined

range of frequencies, in addition to being released over a short time-span. Such a source can be approximated by a sinusoidal signal which is enveloped by a Gaussian rise and decay in amplitude. The model used in LIGO searches for such signals is:

$$h(t) = A \exp\left[\frac{-2(t-t_0)^2\pi^2 f^2}{Q^2}\right] \cos\left[2\pi f(t-t_0)\right], \tag{2.14}$$

for a strain h at time t, with A the amplitude of the signal, t_0 its central time, Q the quality factor of the burst, and f is frequency.

An example of a time-domain sine-Gaussian burst signal($q = 8$, $f = 100\,\mathrm{Hz}$, $A = 1\times10^{-21}$, and $t_0 = 100\,\mathrm{s}$, with linear polarisation) is plotted in the middle column of figure 2.4.1, with the plus- and cross-polarised waveforms both shown.

2.4.4 White noise bursts

Astrophysical processes are unlikely to produce emission at a single frequency, or with a smooth evolution of amplitude, and so searches are normally expected to be sensitive to band-limited white noise bursts, which consist of band-limited uncorrelated noise within a Gaussian amplitude envelope. An example of a time-domain white noise burst (with duration $0.05\,\mathrm{s}$, $f = 1000\,\mathrm{Hz}$, $A = 1\times10^{-21}$, and $t_0 = 100\,\mathrm{s}$, with linear polarisation) is plotted in the right column of figure 2.4.1, with the plus- and cross-polarised time-domain waveforms both shown.

2.4.5 Ringdown-like bursts

Ringdown-like signals, with a sudden rise, and exponential decay in amplitude are expected in the post-merger signal of CBC systems, and in some models of neutron star model excitation [178]. These take the form

$$h(t) = \exp(-t/\tau)\sin(2\pi f t) \tag{2.15}$$

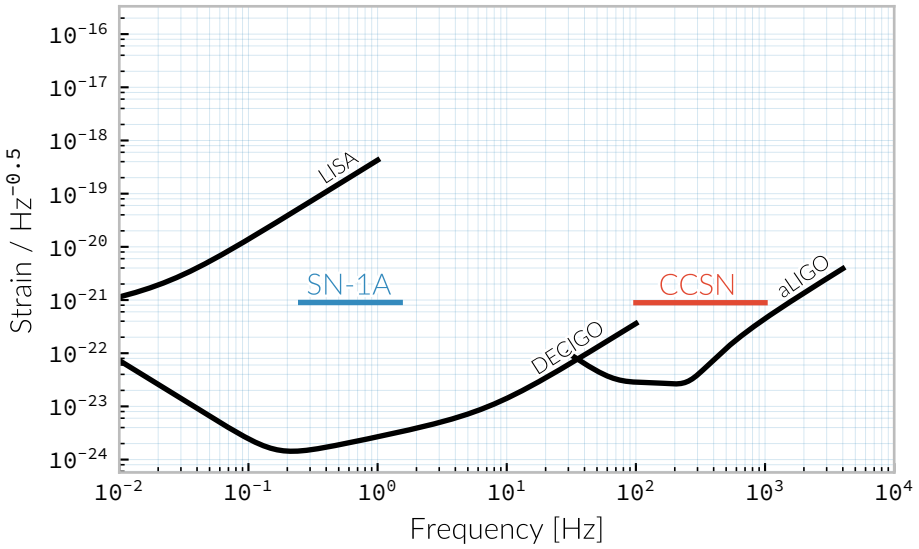


Figure 2.4.2: The frequency ranges and approximate GW amplitudes likely to be produced by core collapse supernovae [179] and Type 1A supernovae [180], at a distance of $10\,\mathrm{kpc}$ from the earth. The noise (sensitivity) curve of the advanced LIGO, DECIGO, and LISA detectors at their design sensitivity is plotted for reference.

for a strain h at time t, given a decay time τ and frequency f.

2.4.6 Core-collapse supernovae

The collapse of a massive star's core is driven by the release of gravitational energy, creating a core-collapse supernova (CCSN). The progenitor stars of CCSNe have zero-age-main-sequence (ZAMS) masses in the range $8\,\mathrm{M}_\odot \leq M \leq 130\,\mathrm{M}_\odot$. Much of this energy is stored as heat in the proto-neutron star (PNS) remnant, around 99% of the released energy is carried-off by neutrinos, around 1% provides the kinetic energy of the explosion, while less than 0.01% of the energy is extracted as electromagnetic and gravitational radiation [181].

When the iron core of a star exceeds the Chandrasekhar mass it becomes unstable, and undergoes gravitational collapse, and is compressed until the

neutron degeneracy pressure is able to halt the collapse. At this point the core becomes stiff, and the inner core rebounds—a phase of the supernova known as "core bounce". The stiff, ultra-dense remnant of the collapse is a PNS, and the rebounding material forms a shock wave. This shock wave is not sufficient, however, to produce the observed explosive phase of CCSNe events, and so must undergo *revival* by some poorly-understood mechanism to lead to the final explosion. GWs are expected to be emitted in a number of periods during the collapse, for example during a rotating collapse, and the core-bounce which follows it; pulsations of the PNS [182]; and anisotropic neutrino emission [183, 184, 185].

In order to predict the gravitational waveforms which would be produced by a CCSN detailed numerical modelling must be completed, with the most modern results from Scheidegger [186], modelling rotating, axisymmetric collapses in three dimensions, and Dimmelmeier [179] in two dimensions. Those from Müller [187] and Ott [188] model neutrino-driven supernovae in three dimensions.

It is possible that nearby core-collapse supernovae could have been detected with the initial LIGO detector [52], although none were. At design sensitivity the three-detector network of Advanced LIGO and Advanced Virgo should be able to detect CCSNe to a distance of $5.5\,\mathrm{kpc}$, but with large uncertainties, in the case of neutrino-driven explosions, while rapidly-rotating core-collapses will be detectable to $50\,\mathrm{kpc}$, the distance to the Large Magellanic Cloud. Extreme emission scenarios may be detectable as far as $0.77\,\mathrm{Mpc}$, the distance to M31 [189]. The characteristic strain spectrum of a CCSN is plotted in figure 2.4.2, alongside the sensitivity curve of both advanced LIGO and two proposed space-based detectors, DECIGO and LISA.

2.4.7 Cosmic strings

Cosmic strings are theorised topological defects which were first postulated by Kibble in 1976 [190]. These are expected to have been produced as a result of phase transitions in the early universe, and carry large quantities of energy. The simplest string models are characterised by the energy density of the string μ, and its tension, which are taken to be equal. The dimensionless quantity $G\mu \sim (T_c/M_{\rm Pl})^2$, with G Newton's gravitational constant, T_c the temperature at the transition, and $M_{\rm Pl}$ the Planck mass, characterises the strength of interactions between strings. For strings produced by the decoupling of the strong force from the electroweak force this quantity has a value on the order of 10^{-6}, so a quantity μ_6 is often defined as a shorthand [191].

Three separate models of cosmic strings have been searched for in data from the advanced LIGO detectors to date, and while no evidence for GW emission from these objects was found, it was possible to place limits both on the parameters of the various models and on the scale of $G\mu$. The LIGO results place a limit of $G\mu$ be less than 4.2×10^{-10} [192], which agrees with, but is surpassed by results from pulsar timing arrays, which find $G\mu$ is less than 5.7×10^{-12} [193].

2.4.8 Parabolic and hyperbolic encounters

Encounters between pairs of black holes, where the two bodies trajectories are affected by the total gravitational field, but where a closed orbit is not formed are expected to be possible in regions of space with a high density of compact objects, for example globular clusters and the centres of galaxies. In the case where the deflection angle of the trajectories is small this process can be considered analogous to Bremsstrahlung processes in electromagnetic radiation production [194, 195], but the emission produc-

tion becomes more complicated as larger deflection angles are considered, and spin is included. Approximate models are available for the waveforms of these encounters in the Bremsstrahlung case, low-velocity cases with arbitrary deflection [196], and head-on collisions [197]. Recent advances have allowed the production of 3.5 PN accurate waveforms for hyperbolic encounters for non-spinning pairs of black holes [198]

2.4.8.1 Encounter waveforms

Recent advances in NR modelling have allowed the production of accurate waveforms for parabolic encounters between spinning black holes, and in this section I present the results of a study I have conducted to consider the detectability of some of these waveforms in current and future detectors. The waveforms for this brief study are taken from the GW driven capture simulations of [199]. An example of one of these waveforms is plotted in figure 2.4.3.

The detectability of the waveforms can be estimated using equation 1.15, taking the Fourier transform of the strain data from the NR simulation to form $\tilde{h}(f)$, and the estimated noise amplitude spectral density (ASD) for each detector. Figure 2.4.5 shows the signal-to-noise ratio (SNR) of the waveform plotted in figure 2.4.3 as a function of the luminosity distance of the source, and the total mass of the system. An SNR of 8 is a standard threshold which a signal must exceed to be considered a detection, and this is plotted with a heavier line. The spectrum of this waveform is plotted in figure 2.4.4, alongside the ASD for advanced LIGO (at design sensitivity), Einstein Telescope (ET), and DECIGO. This plot clearly shows that a signal from a system such as this at a distance of $50\,\mathrm{Mpc}$, a distance which encompasses not only the Local group, but nearby galaxy clusters such as the Virgo group, is above the noise-floor of even advanced LIGO

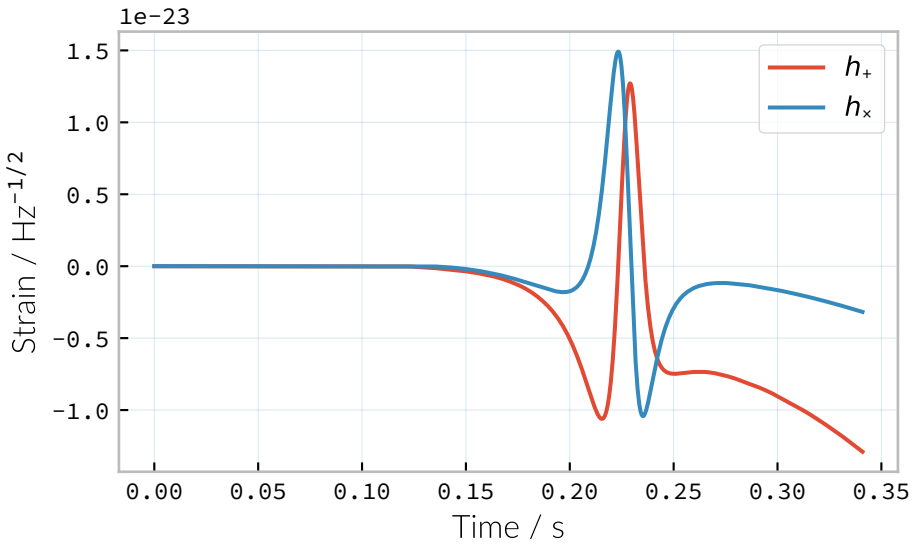


Figure 2.4.3: A parabolic encounter waveform from [199] for a system with mass-ratio $q = 4$, an impact parameter of $5060\,\mathrm{M}_\odot$, and a total mass of $50\,\mathrm{M}_\odot$ at a distance of $50\,\mathrm{Mpc}$. The behaviour of this waveform at times after the main burst of strain indicates that additional high-pass filtering is required when making calculations using the timeseries thanks to the existence of low-frequency artifacts which prevent the waveform returning to zero at large times. The spectrum of this waveform is plotted in figure 2.4.4.

at its design sensitivity. Other systems, which are highly asymmetrical (with a mass-ratio $q = 16$) may be even more promising candidates for detection in advanced LIGO. For example the waveform used to produce figure 2.4.6, with a distance around $175\,\mathrm{Mpc}$; again producing a detectable SNR to distances which encompass nearby clusters of galaxies. This makes these events a promising burst source candidate, however further work is required to establish plausible astrophysical rates for such events in order to determine if these are a likely source of signals for current generation detectors.

Encounters which result in capture may appear similar to high-mass BBH events in the current generation of detectors, since the inspiral por-

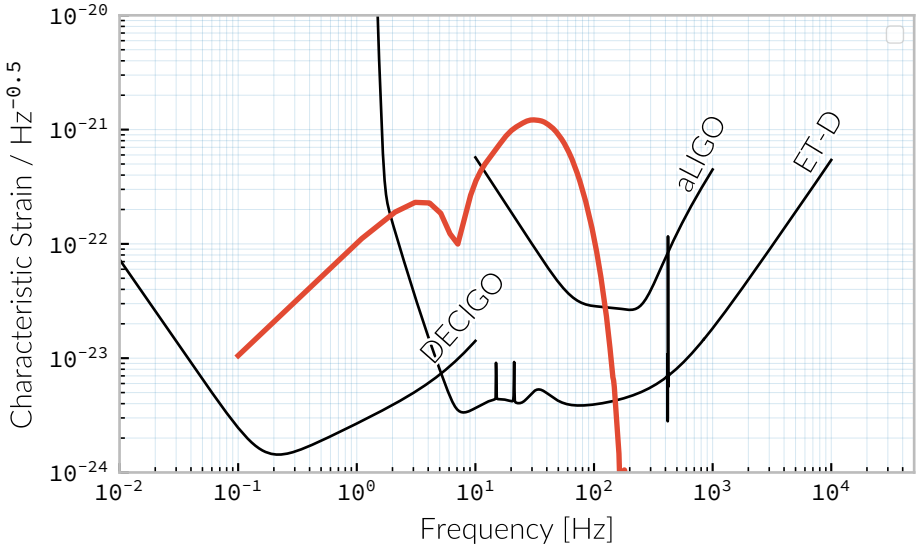


Figure 2.4.4: The spectrum of a parabolic encounter from [199] for a system with mass-ratio $q = 4$, an impact parameter of $5060\,\mathrm{M}_\odot$, and a total mass of $50\,\mathrm{M}_\odot$ at a distance of $50\,\mathrm{Mpc}$ (the same waveform depicted in figure 2.4.3), with the sensitivity curves of LIGO, DECIGO, Einstein Telescope, and LISA for reference. The timeseries representation of this waveform is plotted in figure 2.4.3.

tion of a high-mass system will be predominately at frequencies below which the detectors are sensitive. Encounter waveforms will also lack a lengthy inspiral portion; as a result there is a chance that future GW events may be misclassified, and this eventuality should be considered in waveforms which seem to lack a clear inspiral. These results are principally low-frequency sources, and as such are ideal candidates for both ET, which aims to achieve much greater low-frequency sensitivity than current detectors, but also for decihertz detectors, such as DECIGO.

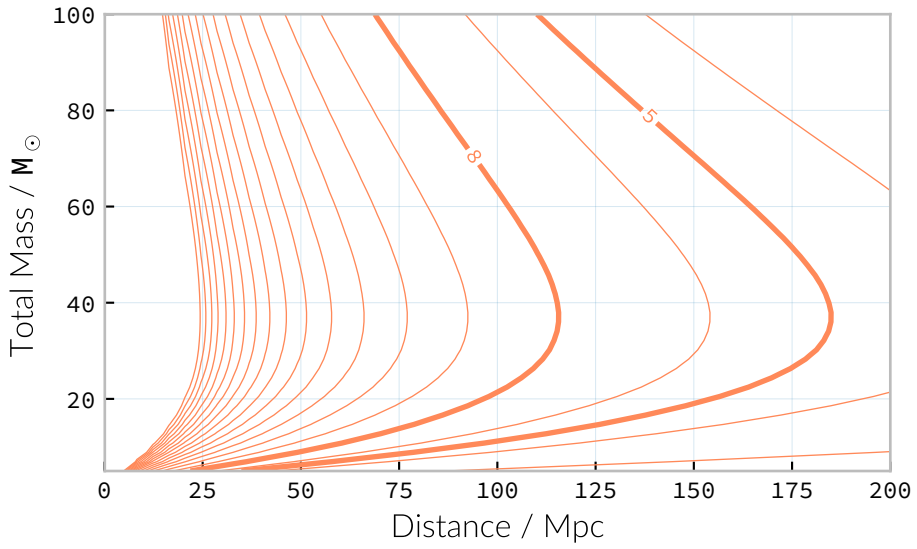


Figure 2.4.5: The SNR of the $q = 4$ hyperbolic encounter waveform plotted in figure 2.4.3 in advanced LIGO at design sensitivity. The heavy line for an SNRs of 8 represents a standard detection threshold used in single detector transient searches.

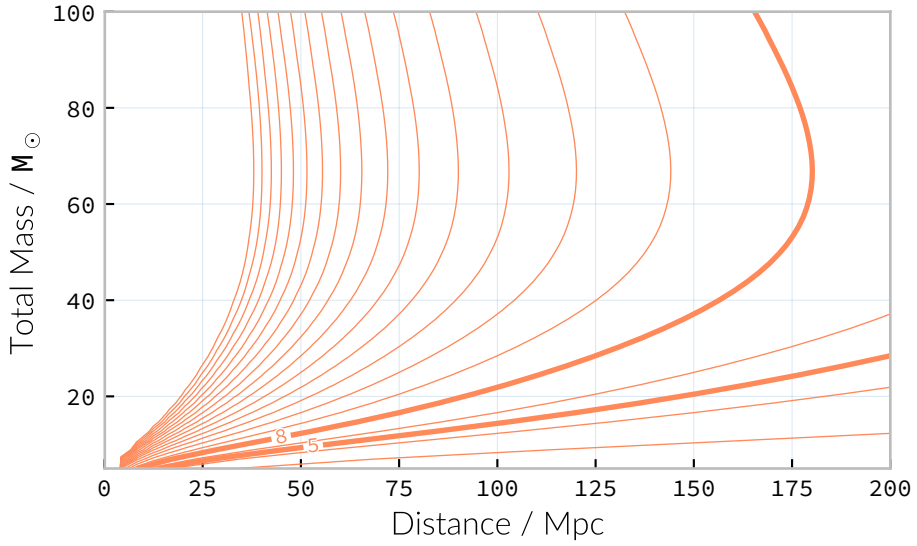


Figure 2.4.6: The SNR of a $q = 16$ hyperbolic encounter waveform in advanced LIGO at design sensitivity. The heavy line for an SNRs of 8 represents a standard detection threshold used in single detector transient searches.

3 Burst searches and mock data challenges

This chapter introduces searches for unmodelled gravitational wave transients ("bursts"), and then discusses the development of the Minke software library which I developed during the course of my PhD, which is used to produce mock data challenges, which are used to measure the sensitivity of burst search algorithms.

Section 3.1 then gives an overview of techniques used to search for these signals in detector data.

Section 3.2 discusses mock data challenges, and how these are used to characterise search algorithms.

Section 3.3 introduces Minke.

Section 3.4 then provides a discussion of the measured sensitivity of a variety of burst search algorithms which were employed on detector data from the first two advanced-era observing runs. This is largely a discussion of the results from [192], which used mock data challenges produced by the Minke library.

In the world of gravitational waves (GWs) there are *known knowns*, astrophysical sources, including binary black holes (BBHs) where the source is well understood, and the GW signal it produces is also well understood. We also know that there are *known unknowns*: astrophysical sources which have been observed electromagnetically, and which are expected to produce GWs, however for these sources the signal they should produce is poorly understood, and poorly modelled. But we know there will be *unknown unknowns*: each development in observational astronomy has been accompanied by unanticipated discoveries. It is unlikely that GWs will be any different, but detecting signals with no prior knowledge of the physics which generates them is a formidable challenge. "Burst" searches are designed to identify signals in detector data which fall into the latter two categories: the known unknowns, such as supernovae, and the unknown unknowns.

3.1 | Burst searches

Burst searches cannot rely on well-known template waveforms in the way that compact binary searches can, and so matched filtering techniques cannot be used. Performing analysis of the signals from the data collected from a gravitational wave detector is complicated by the presence of both stationary and non-stationary noise (see section 1.4 for a discussion of the sources and extent of this noise), and by the signal's convolution with the detector's antenna pattern (see section 1.3.2.5). The final construction of the signal detected by a network of detectors is outlined in figure 3.1.1.

Instead burst searches, similarly to searches for the stochastic background, make use of information gained from correlations between detectors in a network.

There are two approaches to analysing data across a network of detec-

tors.

Coherent analysis which combines the data steams of detectors together into a single stream, with all of the data analysed in the same process. An outline diagram of the operation of a coherent burst search is provided in figure 3.1.2.

Coincident analysis performs a search for signals on each detector's data separately, providing a list of times at which a candidate signal (or trigger) is identified. These are then compared, allowing for suitable time delays corresponding to the wave travel time between detectors in the network, to identify coincident events. An outline diagram of the operation of a coincident burst search is provided in figure 3.1.3.

The coherent method is substantially more difficult to perform, and can require access to greater computational resources than the simpler, faster coincident method. However, the coincident method is generally less sensitive, as a signal which is weakly detected in one detector, but strongly in another may not produce triggers in both analyses, where a coherent analysis would help to identify the more weakly-detected signal.

Initial Laser interferometer gravitational-wave observatory (LIGO) searches were performed both between the detectors constituting the LIGO network (the two 4-km detectors at LIGO Livingston Observatory (LLO) and LIGO Hanford Observatory (LHO), in addition to the 2-km detector at LHO), and between this network and TAMA, GEO600, and AURIGA.

At this stage it will be useful to define a few quantities which are used in search algorithms.

Definition 3.1.1 (Null stream energy)**.** The null stream energy is the minimum amount of energy in whitened detector data which is inconsistent

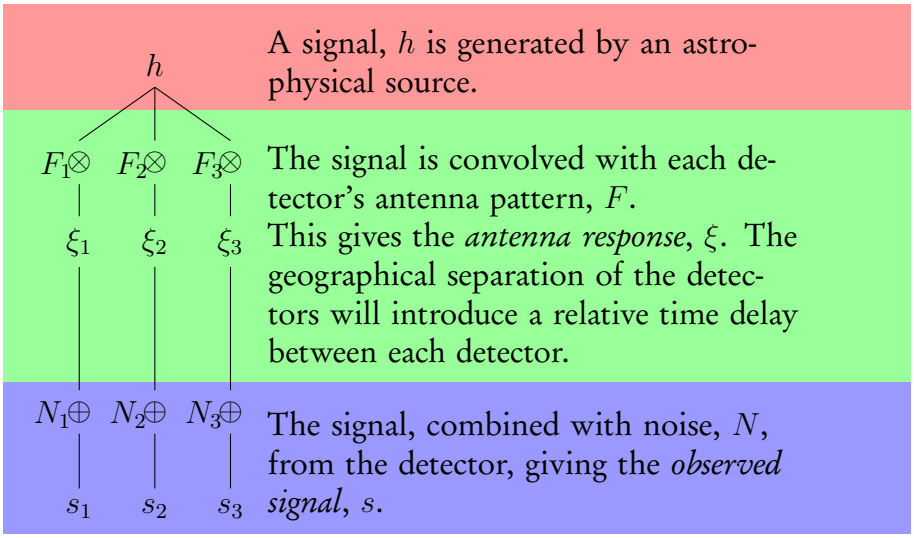


Figure 3.1.1: The construction of the GW signal observed by a network of detectors, from the GW source through to its measurement by a detector.

with a GW signal from a given sky location, across a network of detectors.

Definition 3.1.2 (Excess energy)**.** The excess energy of a single detector measurement is the amount of energy which is not consistent with a noise hypothesis in whitened detector data.

3.1.1 Fundamental search methods

While a large number of algorithms for searching for burst signals exist, most rely on a small number of fundamental techniques to identify burst events in detector data.

Power filter Calculates a weighted spectrogram of the data by splitting the whitened detector data into overlapping chunks. [200]

Wavelet transform These methods uses either a wavelet or a Q-transform [201] approach to produce a time-frequency representation of the measured signal. This is then thresholded, and clusters of outlier pixels

(discrete regions of the time-frequency plane) are identified. By performing a number of Q-transforms with varying q parameters it is possible to estimate the parameters of the detected waveform [202].

Mean filter The mean filter is a time-series approach to burst detection, which searches for excesses in the moving average of the measured signal. The method is most sensitive to bursts which have a similar length to the moving window, so the search must be repeated for a number of different window lengths [177].

3.1.2 Search pipelines

While the fundamental methods described in section 3.1.1 are plausible methods for identifying burst signals in data, they are unable to operate in isolation. This has lead to the need to construct *pipelines* which are capable of pre-processing the detector data, performing searches to produce lists of triggers, estimate the significance of these triggers, and perform parameter estimation on the signals. There are at least four major burst search pipelines in use during the advanced detector runs.

X-Pipeline X-Pipeline is designed to run coherent triggered searches for GW bursts, motivated by the detection of events such as short gamma ray bursts (sGRBs) [203]. The analysis constructs time-frequency spectrograms of the plus and cross strain polarisations, and the null stream, after they have been whitened and time-shifted. Pixels in the spectrograms are then clustered in order to identify significant outliers from the noise. This process is repeated for each location on the sky being searched, with appropriate time-shifts, for each detector's data. Events are vetoed if they have a strong correlation between the coherent energies and incoherent energies, a feature

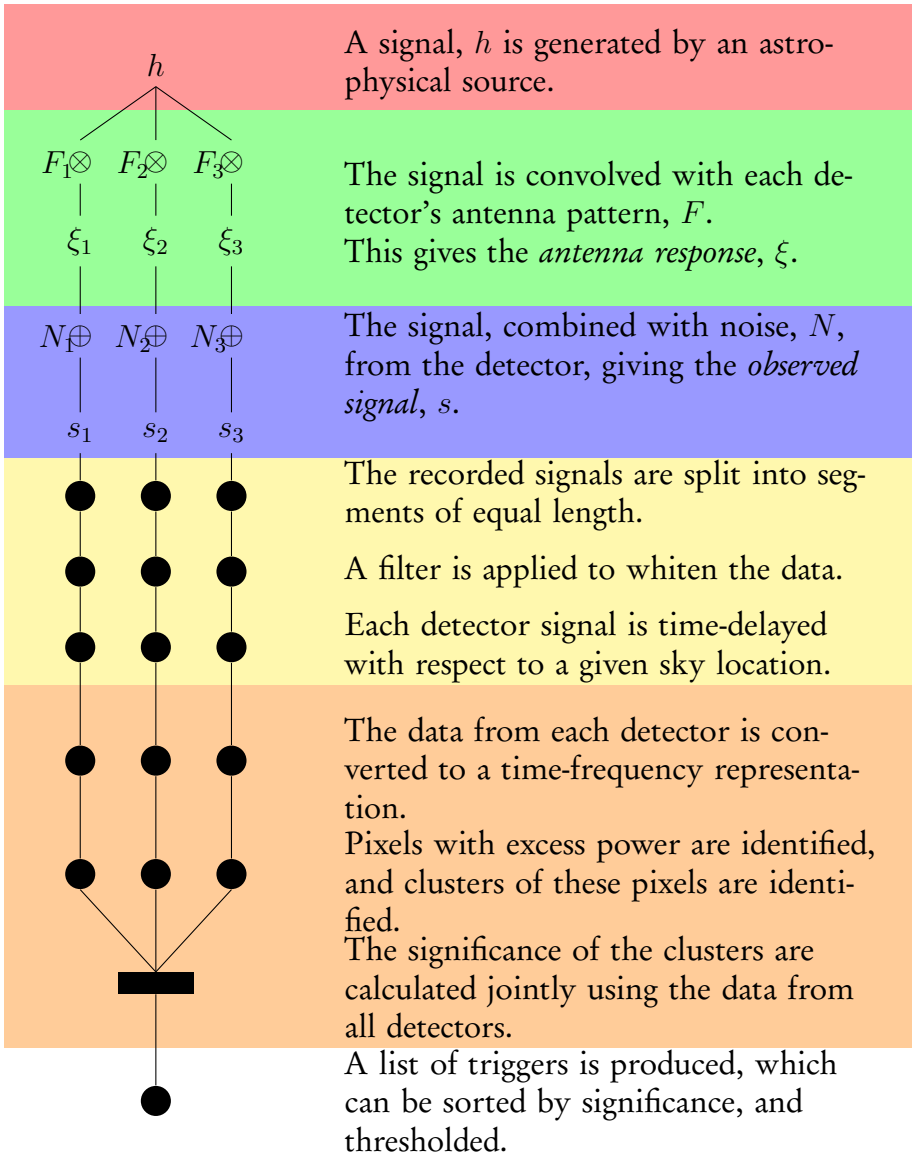


Figure 3.1.2: The principles of a coherent all-sky burst search pipeline.

which indicates a noise glitch rather than a burst signal. The X-Pipeline can be combined with the SphRad pipeline to run all-sky, untargetted searches [204] in the XSphRad configuration.

Coherent WaveBurst The Coherent WaveBurst (cWB) pipeline [205] is a coherent, untriggered burst search method which performs a wavelet transform on blocks of detector data to first produce a time-frequency representation. The wavelet layers are then whitened with a linear prediction error filter, and time-delayed. Correlations and excess-power regions in the time-frequency plane are then clustered to identify coherent triggers, which are then selected by thresholding based on the false alarm probability (FAP).

Omicron / LALInference burst The Omicron LIB (oLIB) pipeline [206] is a coincident all-sky burst search pipeline which relies on the use of Q-transforms to generate time-frequency representations of detector data, in order to identify regions of excess energy. The significance of single-detector triggers are then determined using Bayesian inference to produce a joint detection significance from the network of detectors.

Bayeswave In contrast to other burst search pipelines, Bayeswave (BW) [207] is designed to determine the significance of pre-determined triggers, and does not generate triggers on its own. It takes a direct approach to distinguishing signal transients (bursts) from noise transients (glitches) by directly modelling glitches with Morlet waveforms, and then performing Bayesian model selection to identify a favoured hypothesis: either a noise, glitch, or signal model.

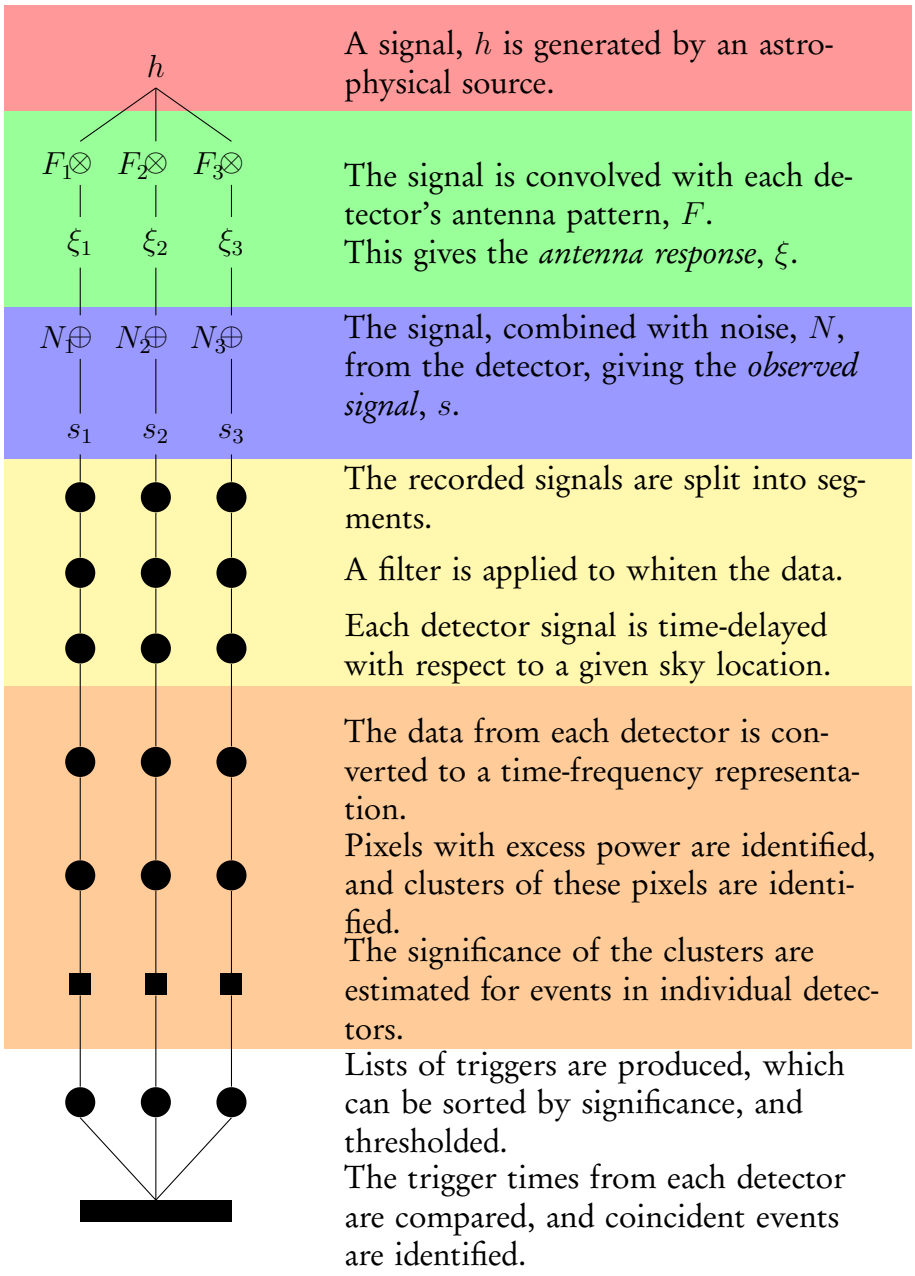


Figure 3.1.3: A typical all-sky coincident burst search pipeline.

3.2 | Mock data challenges and all-sky searches

Given the complexity of modern burst search algorithms, and their need to operate in a regime of non-Gaussian, non-stationary noise it is important to be able to demonstrate the efficacy of an individual pipeline. In addition to this, the nature of the noise in GW detectors makes estimating the sensitivity of each search method difficult, as the time-evolution of the noise power spectral density (PSD) will affect the sensitivity over time. To address these problems pipelines are tested against known signals which are "injected" into the recorded data from the detector.

This process can be performed in two ways, either as a *hardware injection*, where the test mass of the detector is physically actuated (using either an electrostatic drive, or through photon pressure from a laser placed behind the test mass), or as a *software injection*, where the signal waveform is added to the pre-recorded data. The former process has the advantage of testing the performance of the pipeline "end-to-end", as the signal will be present in the analysed data at all times. Indeed, prior to the detection of GW150914 the use of "blind" injections to test the readiness of search algorithms was a standard practice[1]. In contrast software injections can be performed offline, and thus do not run the risk of obfuscating an astrophysical signal. However, as these are added to the recorded data

[1]Indeed, a major undertaking towards the end of the initial detector era had involved the analysis of a signal, identified as a BBH coalescence in Canis Major, which was dubbed the *Big Dog* event. After a thorough end-to-end analysis of the signal was conducted it was revealed to have been such a blind injection. While this exercise demonstrated that the various search algorithms in operation at the time were capable of identifying a signal, it also demonstrated a number of failings in the various components of the software chain which was used to produce and inject this signal into the detector. In fact, the intended signal had been an neutron star / black hole (NSBH) signal, and had not been injected at a location in CMa. A fuller discussion of the event can be found at `https://www.ligo.org/news/blind-injection.php`.

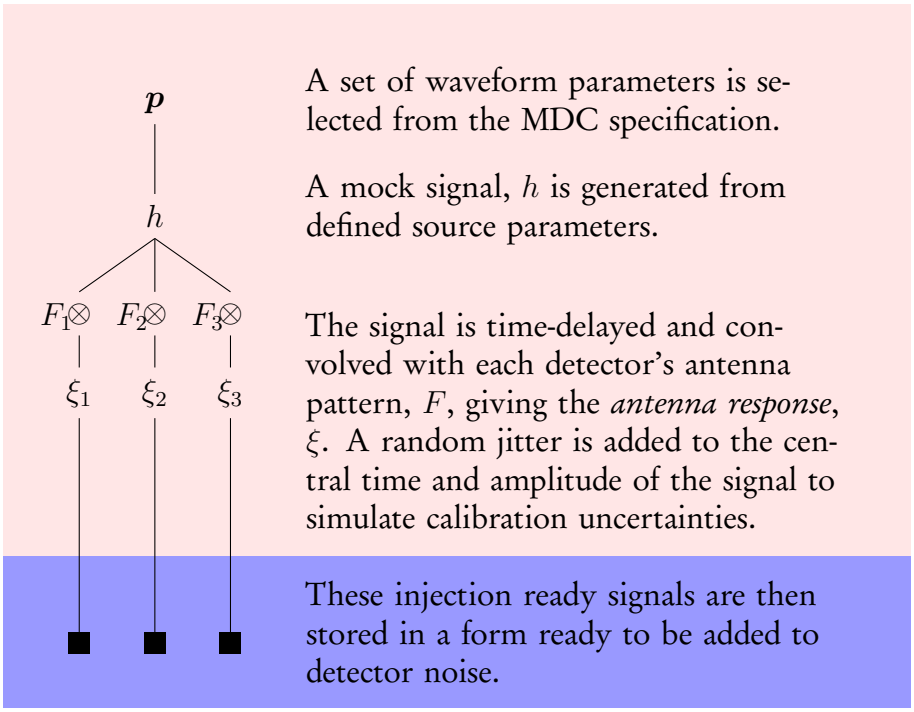


Figure 3.3.1: The process of frame production for a MDC using Minke. For software injections the injection-ready signals are normally stored in GWF frame files, ready to be injected into the signal recorded from the detector. For waveforms to be used for hardware injections the signals are simply written out as ASCII-format text files.

it is necessary to ensure that a consistent set of signals is analysed by all pipelines to provide consistent sensitivity estimates.

At the beginning of the advanced detector era it became clear that the production of these *mock data challenges* required new infrastructure, and the *Minke* project was initiated to handle this.

3.3 | Minke

Minke [208] is a software library[2] implemented in the `python` programming language which is designed to produce large-scale MDCs for burst waveforms. In order to ensure consistency with other GW analyses, and to make use of as much pre-existing, reviewed code as possible, Minke makes use of as much functionality as possible from pre-existing code released as part of the LALSuite [209] library of GW analysis software. While Minke was designed for the primary purpose of MDC production, it also provides a convenient python wrapper to a number of functions from the `LALSimulation` [209] C library, making it a convenient means to generate GW signals, for example for machine learning training data.

Burst searches are used to search for both totally unmodelled sources, and poorly modelled sources (see section 2.4 for a discussion of the various morphologies of these signals). Minke supports a large subset of these waveforms; analytical waveforms, such as Gaussians and sine-Gaussians are generated using `LALSimulation`, while more complicated models can be used (such as numerical relativity (NR) derived supernova waveforms) can be used to generate signals if the waveform is available either as pre-computed strain values in the $+$- and $\times$-polarisations, or decomposed into a spherical harmonic basis.

The process for producing an MDC using Minke is broken into a number of stages (which are also depicted in figure 3.3.1):

1. The distributions of source parameters are specified; these include the probability distributions from which parameters of individual signals are drawn, and the hyperparameters defining those distributions. For

[2]The Minke source code can be obtained from `https://git.ligo.org/daniel-williams/minke`, and is documented at `https://daniel-williams.docs.ligo.org/minke/`. It can also be installed using the `pip` package manager by running `pip install minke`.

example, the injections may be made uniformly across the sky, or a more specific distribution may be chosen.

2. A specification for the signal set is generated as a table of source parameters, with each row specifying a single signal. This table can be stored in XML format for later use.

3. Each signal can then be generated using `LALSimulation`. In the case of analytical waveforms the appropriate generating function is evaluated, and in the case of a pre-computed signal the waveform data is interpolated appropriately for the desired sample rate, and where appropriate, is reconstructed from a spherical harmonic basis.

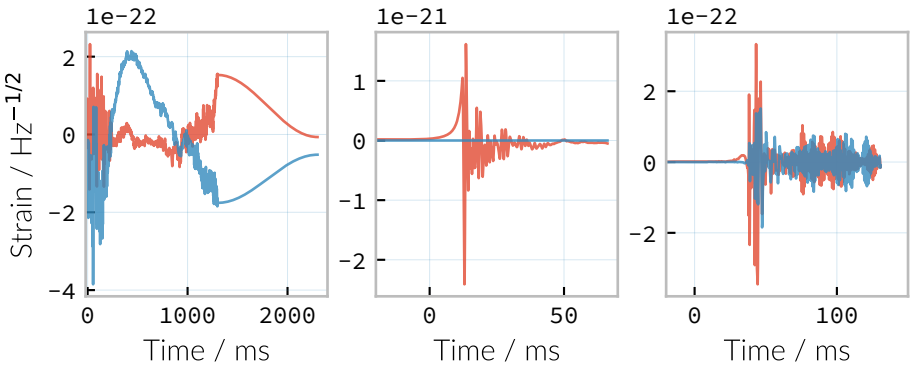
4. The generated signal must then be convolved with the antenna pattern for each detector involved in the analysis, for the appropriate sky position and time, and must have the appropriate time delay applied relative to the geocentre. A random "jitter" may also be applied at this stage to the centre time and amplitude of the signal, to emulate timing and calibration uncertainties in the detector.

5. The signals, which are now ready to be injected, can be stored either in a GWF frame file, or as ASCII data. The former is used for software injections, and the latter for hardware injections.

6. The analysis-ready MDC is constructed by adding the injection signal to the detector data, allowing the sensitivity of the detector to a specific signal morphology to be tested over time, in the presence of real noise and glitches.

When injecting signals which possess GW memory effects (that is, when the strain at the end of the signal does not equal the strain at the beginning) it is desirable to remove this effect. When adding a signal

Figure 3.3.2: Three of the supernova burst "waveforms" which can be simulated and injected by Minke. From left to right, Mueller [210], Dimmelmeier [179], Scheidegger [186] waveforms are depicted in the time domain, with both the plus polarisation (red) and cross polarisation (blue) depicted.



with memory to pre-recorded noise there will be a sharp discontinuity at the end of the signal which will produce artefacts in any frequency-domain representation of the signal. To account for this Minke adds a low-frequency half cosine decay to the end of the signal to present a smooth transition.

Examples of the supernova (SN) waveforms which Minke is capable of injecting are shown in figure 3.3.2. In the Mueller waveform signal [210] in the left column the smooth transition which is added by Minke is clear.

3.4 | O1 and O2 all-sky search sensitivity

The data from the first two observing runs of the advanced LIGO detector were analysed by three all-sky burst search pipelines: cWB, oLIB, and BW. These pipelines, which employ techniques described in section 3.1, did not make detections of non compact binary coalescence (CBC) events during these two observing runs, but were able to place limits on the rate of such events within the sensitive volume of the detector. In order

to determine this sensitive volume MDCs were constructed using Minke (described in the previous section). Signals from each MDC were injected into the strain data collected from each detector every $50\,\mathrm{s}$; these signals had a distribution of root square sum strain (hrss) values, such that the distance distribution of the events, $p(r) = r + 50/r$, for distance r. The resulting hrss distribution will contain both "loud" events, with a high hrss, and "quiet" events with low hrss.

The injection process was repeated for a number of waveform morphologies, listed in table 3.4.1, allowing the pipelines to be assessed across a range of frequencies. The same table indicates the hrss of the weakest signal which was detected with 50% efficiency given a FAR of one per hundred years.

These results indicate that between O1 and O2 there are substantial improvements in the detection efficiency for some signals, with changes to the algorithm allowing some waveforms, such as the $849\,\mathrm{Hz}$ sine-Gaussian, to be detected with much greater efficiency. BW is noticeably the least sensitive pipeline to sine-Gaussian signals during O2, as it struggles to differentiate simple signals which can be represented using a single sine-Gaussian basis from noise.

The GW energy, E_{GW} can be found from the hrss through the relationship

$$E_{\mathrm{GW}} = \frac{\pi^2 c^3}{G} r_0^2 f_0^2 h_0^2, \tag{3.1}$$

for a distance r_0, (fixed) frequency f_0, and hrss h_0.

Each result in table 3.4.1 can then be converted into a representative energy, taking a fiducial distance $r_0 = 10\,\mathrm{kpc}$, and these values are plotted in figure 3.4.1.

The results presented in figure 3.4.1 indicate that the sensitivity of the three burst search pipelines has improved across the range of observed

Waveform parameters			Observing run 1			Observing run 2		
			cWB	oLIB	BW	cWB	oLIB	BW
Gaussian								
	τ/ms							
	0.1		34	NA	NA	8.4	6.2	NA
	2.5		33	7.4	NA	11	5.3	NA
Sine-Gaussian								
f_0/Hz	q							
70	3					4.9	-	NA
70	100		24	NA	NA	6.4	-	NA
153	8.9		1.6	1.7	5.4	1.4	1.3	16
253	100		14	19	NA	3.3	1.1	1.4
554	8.9		2.6	2.7	3.6	1.8	1.5	NA
849	3		27	3.3	5.4	5.5	2.0	17
1304	9					3.3	2.8	-
1615	100		5.5	-	-	3.6	3.3	-
2000	3		8.7	-	-	5.4	5.3	-
2477	8.9		11	-	-	7.5	-	-
3067	3		15	-	-	9.7	-	-
White-noise burst								
f_{low}/Hz	Δf/Hz	τ/ms						
100	100	0.1	2.0		3.0	1.4	3.0	3.0
250	100	0.1	2.2	NA	9.2	1.4	3.8	3.8
750	100	0.1				1.8	3.7	4.2

Table 3.4.1: The hrss values, in units of $10^{-22}\,\text{Hz}^{-1/2}$ at which each pipeline achieves a 50% detection efficiency at a FAR of 1 in 100 yr in both the first and second advanced LIGO observing runs, thus a lower value implies a more sensitive search. A value of NA indicates that 50% efficiency could not be achieved, while "-" indicates that the search was not run on a given morphology, as the characteristic frequency did not meet the criteria of the search. The data in this table is derived from tables presented in [211] (O1) and [192] (O2).

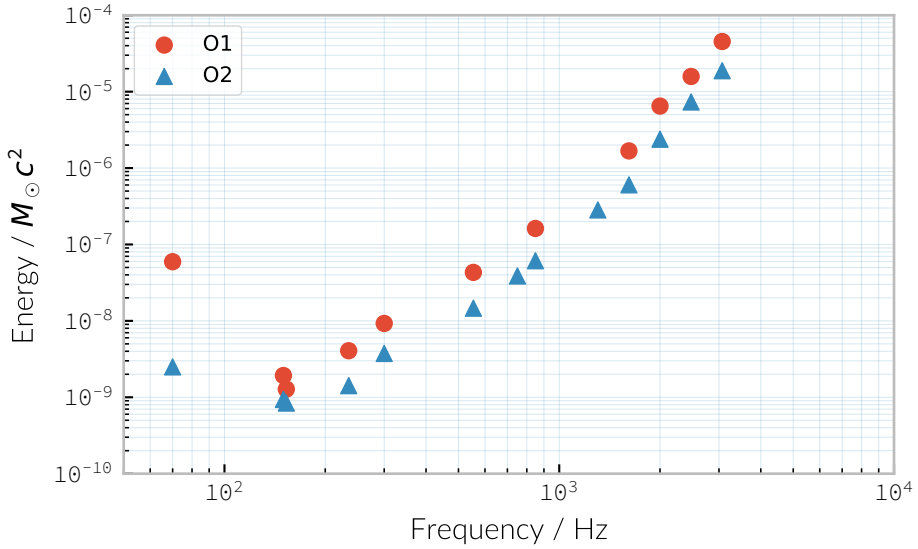


Figure 3.4.1: The emitted GW energy, in units of solar masses, which correspond to a 50% detection efficiency at a given frequency, for a FAR of 1 event per hundred years, and at a fiducial distance of 10 kpc. At each frequency the marker corresponds to the best performance from the three search pipelines. The data to produce this plot were derived from the values in table 3.4.1, for the sine-Gaussian and white-noise burst waveforms.

frequencies between the first two observing runs, but the data in table 3.4.1 indicate that the improvements are not uniform across all searches, with some searches altering their configuration in such a way that they gain sensitivity in some frequencies while reducing it in others. The greatest sensitivity of the network of GW detectors to burst signals is therefore achieved by considering the results from all the available search pipelines.

3.5 | Summary

The development of Minke as a means of testing the sensitivity of all of the burst search pipelines running on data from the advanced LIGO and advanced Virgo detectors has allowed the sensitivity of each pipeline to be compared on fair and equal terms using mock data challenges.

While this is a useful exercise for the purposes of algorithm development, it is also important to properly understand the sensitivity of the detectors to unmodelled signals, as this allows limits to be placed on the rate at which astrophysical events producing GW bursts occur in the local universe, even in the case of non-detection. The ability of the various algorithms to detect signals in mock data challenges, detailed in table 3.4.1 demonstrates why it is necessary that each search pipeline is tested, as each pipeline's sensitivity is frequency-dependent. Understanding the behaviour of each pipeline allows the results from the appropriate pipeline to be used when calculating event rates based across each region of the observed spectrum.

Previously to my development of Minke these tests were often conducted in an ad-hoc manner, and the process of producing an MDC was not standardised or automated. The move to standardised, automated MDC production should allow the continued comparable testing of search pipelines throughout future observing runs.

4 Bayesian inference

This chapter presents an overview and review of statistical probability from a Bayesian perspective, and serves mostly to introduce concepts which will be relied upon in later chapters. While a few examples have been created by adapting classic examples (for example the example of structured distributions in section 4.5 is very nearly the classical "sprinkler" example, with small modifications to be more astrophysically relevant) the material is not novel.

Section 4.1 introduces probability from an axiomatic basis, and section 4.2 introduces concepts from information theory. Section 4.3 discusses how prior information can be treated within a Bayesian framework. Methods for dealing with complicated data using feature spaces are introduced in section 4.4, and structured probability distributions are discussed in section 4.5. The process of statistical inference is introduced in section 4.6, and stochastic processes are introduced in 4.7. Approximate inference techniques are introduced in section 4.8, and hierarchical modelling, to which they are highly applicable, is discussed in section 4.9.

4.1 | Probability

In this work I consider "probability" as a measure of evidential support for a given outcome or event (this is the so-called *quasi-logical* interpretation). For example, a coin, when tossed, can take one of two states when it lands, which I call *heads* or *tails*; I denote these two states H and T. Assuming that I have no knowledge of a reason why the coin would fall one way rather than the other, I am forced to conclude that I have no more evidence to suggest that the coin will fall heads, compared to it falling tails.

With this philosophy, it is possible to construct an axiomatic mathematical model of probability. Here I will follow the approach taken by Kolmogorov [212], approaching from set theory principles. By making some additional demands on this model, one of which is that an event which is considered "certain" should have a probability of 1, it is possible to maintain consistency with boolean logic. Likewise, if an event is certain not to occur, it is assigned a probability 0.

In order to make future discussions more concise I define two concepts related to the configuration of a probabilistic system: *sample space* and *state*.

Definition 4.1.1 (Sample Space)**.** If a variable can take on a number of different values, then the set of all its possible values is called the sample space. In any given problem, the sample space composes the universe set, and is often denoted Ω.

The sample space is analogous to the *configuration space* of a (classical) physical system, or the *state space* of a quantum mechanical one.

Definition 4.1.2 (State)**.** A state is any subset of zero or more elements from the sample space. A state containing one element is a simple state,

while one containing more than one element is a compound state. The states form a σ-algebra on the sample space.

Since the configuration of a probabilistic system is often time-dependent it is often natural to refer to a given state as an *event*.

Kolmogrov then postulates three axioms for probability [213].

Definition 4.1.3 (Probability)**.** Let x be some variable which is capable of having a state $E \in \Omega$ for Ω the *sample space* of the variable. Probability P is a mapping $P : E \to [0, 1]$ which assigns a real value between 0 and 1 to every $E \in \Omega$. We place three constraints on the form of this mapping:

1. For every $E \in \Omega$, $P(E) \geq 0$.

2. $P(\Omega) = \sum_{E \in \Omega} P(E) = 1$.

3. For two states, $A, B \in \Omega$ which are disjoint, such that $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

The first axiom ensures that all probabilities are positive, and that if a state exists within the sample space it has a probability greater than zero. It follows that states which do not exist (i.e. events which cannot occur) have zero probability.

The second axiom fulfils the requirement that the total probability of all possible states should be 1, and that the probability of the system having taking one state from the sample space (i.e. of one of the possible events occurring) is 1.

Finally, the third axiom defines how the probability of subsets within the sample space should be calculated. Provided a set of states are *mutually exclusive* (i.e. that they are disjoint in the sample space), the probability of the states together is equal to the sum of their individual probabilities.

Alternative axioms for probability also exist which do not attempt to define the theory so strongly in terms of measure theory, for example, *Cox's axioms* [214].

The basic logical operations can be extended to probabilities; the equivalent of the **AND** operation becomes the probability of the conjunction of two subsets of the sample space:

Definition 4.1.4 (Joint probability)**.** Given two states, $A, B \in \Omega$, the probability of both states, $P(A \cap B)$ is termed their "joint probability". In the case that these states are independent it is computed as

$$P(A \cap B) = P(A)P(B).$$

Equally, the **OR** operation becomes the probability of the union of subsets of sample space:

Definition 4.1.5. Given two states $A, B \in \Omega$, the probability of either A or B is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

In the case that of two events which occur with some dependence between them, we can form a "conditional probability", for example, if there can be no smoke without fire, then the probability of smoke can be conditional on the probability of fire.

Definition 4.1.6 (Conditional probability)**.** Given two events, $A, B \in \Omega$, then the probability of A *given* B is

$$P(A|B) = \frac{P(A, B)}{P(B)}.$$

If $P(B) = 0$ then $P(A)$ is undefined.

Given that $P(A, B) = P(B, A)$, we have $P(A, B) = P(B, A) = P(B|A)P(A)$, which leads us to a powerful result in probability: **Bayes Theorem** [215].

Theorem 4.1.1 (Bayes Theorem)**.** Given two events, A and B, we may represent the probability of A given B in terms of the probability of B given A:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \tag{4.1}$$

A useful corollary in the case of two independent states A, B (i.e. states which are disjoint in the sample space),

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{(P(A)P(B))}{P(B)} = P(A).$$

There may also be situations where two variables become independent if the state of a third variable is known, providing conditional independence.

Definition 4.1.7 (Conditional independence)**.** Two states, A, B are said to be conditionally independent given a third state, C, if

$$P(A, B|C) = P(A|C)P(B|C).$$

We can denote conditional independence as $A \perp\!\!\!\perp B \mid C$.

From here on I will start to substitute the concept of a state or event for a variable which represents that state, so the notation $P(x)$ will represent the probability of a variable state x. Since a variable can represent a set of potential states, we can introduce a function which maps from the variable to the probability.

In the case of a discrete sample space this function is the probability mass function.

Definition 4.1.8 (Probability mass function)**.** For a discrete variable x, the probability mass function, p, of the variable is the mapping $p(x) = P(X = x)$

In the case of a continuous sample space the mapping p is known as a probability density function (PDF), which is defined

Definition 4.1.9 (Probability density function)**.** For a continuous variable x, the probability density function p of the variable is the mapping p_X such that the probability of a state between a and b is

$$P(a \leq X \leq b) = \int_a^b p_X(x)\,\mathrm{d}x \tag{4.2}$$

It is normal to use the short-hand notation $p(x)$ for the probability of a value x to represent $\int_{-\epsilon}^{\epsilon} p(x)\,\mathrm{d}x$ for a small value of ϵ.

4.2 | Information

Understanding how informative an random variable, X is can provide insight into how well observations of that variable will inform our knowledge of the probability distribution from which it is drawn.

Definition 4.2.1 (Fisher information)**.** Given a PDF, p, for a random variable, X, which is parameterised by a variable θ, the *score*, V of the PDF is defined

$$V(\theta, X) = \frac{\partial}{\partial X} \log p(X, \theta). \tag{4.3}$$

The variance of the score is the *Fisher information* of the distribution:

$$I(\theta, X) = \mathbb{E}(V^2|\theta) = \int V^2 p(X, \theta)\,\mathrm{d}x. \tag{4.4}$$

Knowledge of the Fisher information for a given distribution is particularly valuable in selecting an *uninformative prior* (see section 4.3.1) when designing a Bayesian analysis, where it can be valuable for the prior probability distribution to contribute no information to the inference.

Definition 4.2.2 (Shannon information)**.** Given a PDF, p, for a random variable X the *Shannon information content* of a given value x of X is defined as

$$h(x) = \log_2 p^{-1}(x) \tag{4.5}$$

where the information is measured in *bits* (assuming that a base-2 logarithm is used; if the natural logarithm is used the units are *nats*, and the base-10 gives rise to the *dit*).

Definition 4.2.3 (Entropy)**.** The entropy of a random variable X with a PDF, p is the average Shannon information of the random variable across all its possible values:

$$H(X) = \int p(X)h(X)\,\mathrm{d}X \tag{4.6}$$

taking $0 \log(1/0) \equiv 0$.

4.2.1 Comparing probability distributions

The information difference between two probability distributions, or indeed the information gain of one relative to another can be an important metric when producing inferential models.

Definition 4.2.4 (Kullback-Lieblier Divergence)**.** For two probability distributions, P and Q the Kullback-Liebler Divergence characterises the relative information content of the two, and is defined as

$$D_{\mathrm{KL}}(P, Q) = \int_{-\infty}^{\infty} \log \left[\frac{p(x)}{q(x)}\right] p(x)\,\mathrm{d}x \tag{4.7}$$

A related metric, the Shannon-Jensen divergence is symmetric and always finite.

Definition 4.2.5 (Shannon-Jensen Divergence)**.** For two probability distributions, P and Q the Shannon-Jensen Divergence characterises the relative information content of the two, and is defined as

$$D_{\text{SJ}}(P, Q) = \frac{1}{2} D_{\text{KL}}(P, Q) + \frac{1}{2} D_{\text{KL}}(Q, P) \tag{4.8}$$

4.3 | Prior knowledge

The *prior* probability distribution is perhaps the characterising feature of the Bayesian approach to statistics, whereby the state of belief prior to any observation being made is encoded in a probability distribution. Bayes Theorem allows the *updating* of our state of belief, with the prior distribution being updated by data collected from observation or experiment.

4.3.1 The least informative priors

While the ability to incorporate prior knowledge into an inference is valuable, there are clearly times when we have *no* prior knowledge of a situation. In these situations we must turn to *least informative* priors, which place the same probability on any possible event in the sample space. The simplest approach to constructing such a prior is through the *principle of indifference*, whereby equal probability is assigned to every possible state. For example, if we wished to conduct an experiment to determine the fairness of a 20-sided die, but had no prior knowledge to assume that one side was more likely to be rolled (which is the desirable state for a fair die) then we would assume each side had a probability of $1/20$ of being rolled. In a continuous system such an arrangement is represented as a uniform distribution. Such an approach must be taken with care, however.

Consider the situation in which cube is hidden behind a curtain. We are told that each edge of the cube is between 3 and 5 metres long. We have no further information to indicate which length is most likely, so assign uniform probability to each possibility. The mid-point of this uniform distribution is then $4\,\mathrm{m}$, so we might conclude that to be the most likely length of each side, giving a cube with $16\,\mathrm{m}^2$ faces, and a volume of $64\,\mathrm{m}^3$. We are then told that the surface area of each face is between $5\,\mathrm{m}^2$ and $25\,\mathrm{m}^2$. Making similar assumptions we'd reach the conclusion that the surface area of each face was $15\,\mathrm{m}^2$. This is clearly in tension with our estimate from the edge lengths; clearly the choice of a uniform prior in one set of variables implies a non-uniform one in another.

It is therefore desirable to work with a prior distribution which will vary appropriately under a change of variables [1]; such a prior is known as a *Jeffreys Prior*. A *Jeffreys Prior* which will be invariant under reparameterisation of parameters $\boldsymbol{\theta}$ can be determined from the Fisher information, I:

$$p(\boldsymbol{\theta}) = \sqrt{\det I(\boldsymbol{\theta})} \tag{4.9}$$

4.4 | Feature spaces and Kernels

A feature map is a projection from a lower-dimensional data space to a higher-dimensional one, which can be represented by a mapping, ϕ.

Definition 4.4.1. For a D-dimensional vector $\boldsymbol{x}$, a feature map, $\phi : \mathbb{R}^D \to \mathbb{R}^N$ is a mapping which projects $\boldsymbol{x}$ into an N-dimensional space, the *feature space*.

[1] It is worth noting that in probability and statistics this property is known as *invariance*, but in other areas of mathematics and physics is more likely to be called *covariance*, for example in general relativity.

This can be a valuable technique in statistical regression and classification, where data may become linearly separable in a higher dimensional space, or can be described by a simpler function than in the original data space. An example of such a mapping is $\phi : \mathbb{R} \to \mathbb{R}^3, \quad \phi(x) = (1, x, x^2)^\intercal$, (where $\cdot^\intercal$ is the transpose operator) which can be used to implement quadratic regression, as

$$f(\boldsymbol{x}) = w_0 + w_1\boldsymbol{x} + w_2\boldsymbol{x} = \phi(\boldsymbol{x})^\intercal \cdot \boldsymbol{w} \tag{4.10}$$

which remains linear (and therefore analytically solvable) provided ϕ is independent of $\boldsymbol{w}$.

Once data is mapped from the data space into the feature space it is desirable to have some notion of distance between the features (which we might interpret as the *similarity* between pairs of data). We define a function which computes such a quantity as a *kernel*:

Definition 4.4.2. For all variables x and x' in the input space, $\mathcal{X}$ of a probability distribution, a mapping $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a kernel function.

If the kernel function can be written in the form of a dot-product between two *feature maps*, $\phi : \mathcal{X} \to \mathcal{V}$,

$$k(x, x') = \langle \phi(x), \phi(x') \rangle v,$$

for $\mathcal{V}$ some inner product space, then we can perform the "kernel trick", allowing us to define the kernel in terms of the inner products within the data, without resorting to an external coordinate system.

4.5 | Structured probability distributions

A complicated joint probability distribution can often be factorised into lower-dimensional factor distributions if there are conditional indepen-

dences within the model which that distribution describes. For example,

$$p(a, b, c) = p(a|b, c)p(b, c) = p(a|b, c)p(b|c)p(c).$$

We can then represent these factorisations in the form of a directed graph, with

$$c \to b \to a$$

representing $p(a, b, c)$. In such a graph we use the direction of an arrow to imply a conditional relationship. When expressed in this form we can call the probability distribution a belief network, or a graphical model.

As a concrete (if rather naive) example, consider a situation in which observations are made continuously over the whole sky with two detectors. One is sensitive to gravitational wave (GW) emission, and the other to gamma ray emission. An observing program is established to analyse transient signals detected with one or both of these telescopes, with the belief that GW bursts can be produced by either a binary neutron star (BNS) coalescence, or a binary black hole (BBH) coalescence.

A simple model is constructed which contains four variables

1. $\Gamma \in \{0, 1\}$ which takes the value 1 iff a short gamma ray burst (sGRB) is detected,

2. $G \in \{0, 1\}$ which takes the value 1 iff a GW burst is detected,

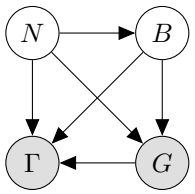
3. $B \in \{0, 1\}$ which takes the value 1 iff a BBH coalescence has occurred, and

4. $N \in \{0, 1\}$ which takes the value 1 iff a BNS coalescence has occurred.

The joint probability distribution of this model is then $p(\Gamma, G, B, N)$, however we can break this down into a structured form by applying the definition of conditional probability (definition 4.1.6),

$$p(\Gamma, G, B, N) = p(\Gamma|G, B, N)p(G, B, N) \tag{4.11a}$$
$$= p(\Gamma|G, B, N)p(G|B, N)p(B, N) \tag{4.11b}$$
$$= p(\Gamma|G, B, N)p(G|B, N)p(B|N)p(N) \tag{4.11c}$$

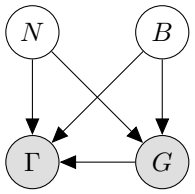
We can represent this model as a graph



Our observers have access to a number of up to date astrophysical theories which they can use to further develop the model; these place *conditional independence* constraints on the model.

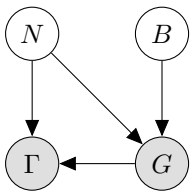
- BBH coalescences and BNS coalescences are independent (one does not cause the other)

This statement implies that $p(B|N) = p(B)$, and $p(N|B) = p(N)$, which we can represent in the graphical form of the model by removing the edge connecting B and N.



- A BBH coalescence does not produce any electromagnetic emission (and therefore cannot produce a sGRB)

This statement implies that $p(\Gamma|B) = p(\Gamma)$, which can be represented in the graphical form of the model by removing the edge connecting Γ and B.



These two constraints considerably simplify the model, and we are now left with the distribution in the form

$$p(\Gamma, G, B, N) = p(\Gamma|N, G)p(G|N, B)p(B)p(N), \tag{4.12}$$

which is easily interpreted from the graphical form of the model, but could have been tedious to derive algebraically.

We can define a belief network more generally as follows.

Definition 4.5.1. A belief network is a probability distribution of the form

$$p(x_1, \ldots, x_N) = \prod_{i=1}^{N} p(x_i|\mathrm{pa}(x_i)), \tag{4.13}$$

where $\mathrm{pa}(x)$ represents the parental set of the variable x; that is, the set of all variables in the graph which have a directed edge ending at x, or the set of all variables on which x is directly conditional.

4.5.1 Equivalence of graphical models

An important caveat with the use of graphical models is that two graphically distinct models may be mathematically equivalent. The reason for this becomes clear when considering the procedure used to factorise the probability distribution starting at equation 4.11a. If we had chosen to re-arrange the variables such that the joint distribution was $p(N, B, G, \Gamma)$

we would have been left with a factorised distribution in which the arrows of the graph pointed in opposite directions, yet this is clearly still the same probability distribution, since probabilities are commutative. To overcome this problem we need to have a definition of equivalence in the graph. A suitable definition is that of *Markov equivalence* [216]:

Definition 4.5.2 (Markov equivalence)**.** Two graphs are Markov equivalent if they both represent the same set of conditional independence statements.

Clearly some method to determine this graphically is warranted. To do so it is helpful to define a (rather judgmentally-named) property:

Definition 4.5.3 (Immorality)**.** Consider three nodes, A, B, and C in a directed acyclic graph (DAG). If C is a child of both A and B, but A and B are not directly connected, then the configuration $A \rightarrow C \leftarrow B$ is denoted an immorality.

In order to determine Markov equivalence we remove all of the directionality from the edges of the graph, producing the skeleton graph. Two graphs are Markov equivalent if they share the same skeleton, and if they share the same set of immoralities.

4.6 | Inference

In section 4.5 I introduced a probabilistic model which consisted of the joint probability of all of the model parameters. Taking the example of joint GW and gamma ray observations, if we know the probability that at any given time there will be a BNS event, we can infer the probability that a sGRB and a GW burst will occur. A model of this form is often considered a "forward model", in that it predicts the probability of an observable, and calculation through the graph follows the arrows. While

such forward models are of considerable utility when attempting to make predictions about unknown variables, often with pre-existing data, they are unable to answer a question such as "given that I have seen a GW, but no sGRB, what is the probability that I have observed a BBH event?". In order to answer such a question we must traverse the graphical model *backwards*, against the direction of the arrows. This process is known as *inference*.

In order to produce the *reverse model* we can turn to Bayes Theorem (theorem 4.1.1). This allows us to derive an expression for $p(B = 1|G = 1, \Gamma = 0)$, that is, the probability that we observe a BBH given that we've observed a GW but no sGRB.

$$p(B = 1|G = 1, \Gamma = 0) = \frac{p(B = 1, G = 1, \Gamma = 0)}{p(G = 1, \Gamma = 0)} \tag{4.14}$$

$$= \frac{\int_N p(B = 1, G = 1, \Gamma = 0, N)}{\int_{B,N} p(G = 1, \Gamma = 0, B, N)} \tag{4.15}$$

$$= \frac{\int_N p(\Gamma = 0|G = 1, B = 1, N)p(G = 1|B = 1, N)p(B = 1|N)p(N)}{\int_{B,N} p(\Gamma = 0|G = 1, B, N)p(G = 1|B, N)p(B|N)p(N)} \tag{4.16}$$

$$= \frac{\int_N p(\Gamma = 0|G = 1, B = 1, N)p(G = 1|B = 1, N)p(B = 1|N)p(N)}{\int_{B,N} p(\Gamma = 0|G = 1, B, N)p(G = 1|B, N)p(N)} \tag{4.17}$$

the probability $p(B = 1|G = 1, \Gamma = 0)$ is called the *posterior probability of* B.

Inference which is based on Bayes Theorem, is a method of statistical inference which is well-suited to situations where a body of evidence grows over time, with new results updating previous understanding of some phenomenon, and as such is well suited to the analysis of experimental data. It is well suited to the analysis of GW data, where measurements are frequently made at different sensitivities during different observing runs.

If we have some hypothesis, some parameters of the hypothesis, I (also called hyperparameters), and some experimental data, we can determine the probability of the hypothesis via

$$p(\text{hypothesis}|\text{data}, I) \propto p(\text{data}|\text{hypothesis}) \times p(\text{hypothesis}, I) \quad (4.18)$$

where $p(\text{data}|\text{hypothesis})$ represents the likelihood; the probability that a given datum would be observed given the hypothesis, and $p(\text{hypothesis}|I)$ represents the *prior* probability, which represents the understanding of the probability of the hypothesis before the experiment was conducted. $p(\text{hypothesis}|\text{data}, I)$ is the *posterior* probability of the hypothesis [217].

Bayesian inference can then be used as a powerful method for *model selection*, where the posterior probabilities of two competing models are compared, with a greater posterior probability indicating greater support for a given model.

4.7 | Stochastic processes

A stochastic process is some collection of random variables which can be indexed by a set, the *index set*. When a stochastic process is used to describe a physical system the indexing set is often taken to be time (represented as either a real or natural number), for example for Brownian motion. Each random variable takes values from its own sample space, Ω. Since each random variable will have a different value each time the process is evaluated, the value of the process as a whole, across all indices, will be different each time. An individual draw from such a process is a *realisation*, or a sample function.

A stochastic process is represented as the set $\{X(t)|t \in \mathsf{T}\}$ for $X(t)$ the random variable drawn indexed by the value t from the index set T.

A simple example of a stochastic process is the **Bernoulli process**, in which each random variable is the result of a Bernoulli test, for example, flipping a (potentially biased) coin. In such a process each $X(t) \in \{0, 1\}$, and $P(X(t) = 1) = p$, with p taking the same value for all t. Because each Bernoulli trial is independent, and all of the trials are equally distributed, the process is independent and identically distributed (iid).

The **Poisson process** extends the concept of a Bernoulli process to the continuous case. Where the Bernoulli process models a discrete state of a system at some given index, the Poisson process models the number of times the system has taken that state in the interval between two indices.

A **Markov process** can be either a discrete or continuous stochastic process where the probability of moving to the next state depends only on the current state of the process, and none of the previous ones. These processes are of considerable importance in Bayesian statistics thanks to their use in various sampling algorithms.

4.8 | Approximate inference methods

In many problems the posterior probability distribution which we need to evaluate will not be analytical. As a result identifying regions of the distribution where the probabilities are large (therefore the areas of interest within the distribution) is likely to require evaluating the function over its entire parameter space, which may be large. This problem is further complicated if the distribution is multi-modal, or contains narrow peaks which may be difficult to find. Further, the evidence term for the posterior is not normally known. The combination of these issues for many distributions makes drawing samples from an arbitrary posterior probability distribution difficult.

For inference, we have two problems which must be solved: how to

generate independent samples from a given probability distribution, and how to estimate the expectation of functions under the distribution.

If we are able to solve the first problem the second can be estimated by using R random samples, $\{\boldsymbol{x}_r | r \in 1, \ldots, R\}$, drawn from the distribution, giving an estimator for the expectation, $\hat{\mathbb{E}}(\phi)$ for the function ϕ,

$$\hat{\mathbb{E}}(\phi) = \frac{1}{R} \sum_r \phi(\boldsymbol{x}_r) \tag{4.19}$$

Given that evaluating a continuous system at every location in its state space is not possible we need a means of producing samples from the distribution which are representative of the distribution. A straightforward approach is to uniformly sample the state space (one strategy to do this would be to devise a grid and take samples at each grid point), however such an approach will work only for the simplest distributions (see chapters 4 and 29 of [218] for a detailed information theoretic discussion on this).

If sampling from the distribution is difficult, but evaluating it at a specific location in its parameter space is possible, a number of sampling methods are possible. The simplest of these, *importance sampling*, and *rejection sampling* rely on sampling from a tractable distribution, such as a Gaussian distribution, and then correcting the samples in some way based on the evaluation of the target distribution.

With *importance sampling*, rather than sampling from the complicated distribution, P, (the *target distribution*), we instead sample from a distribution, Q, which we do know how to sample from, such as a normal or a uniform distribution (see figure 4.8.1 for a cartoon illustrating this arrangement). Since we do not necessarily know the normalisation of P or Q we can instead sample and evaluate within a scalar multiple, Z, such that $ZP^*(x) = P(x)$. We then draw the samples $\{\boldsymbol{x}_r | r \in 1, \ldots, R\}$ from Q, and evaluate $Q(x)$ and $P(x)$ for each sample. In regions where $Q(x)$ is greater than $P(x)$ the samples will over-represent $P(x)$ (and vice

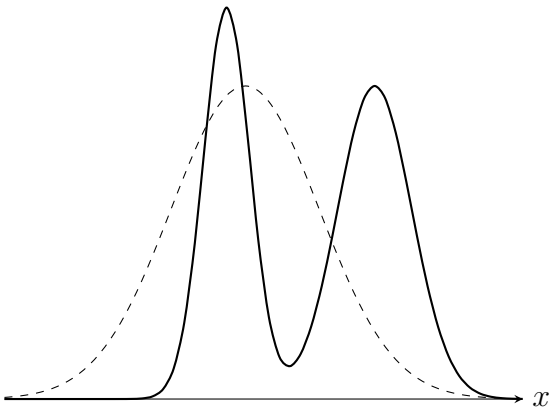


Figure 4.8.1: In importance sampling the arbitrarily complicated distribution, $P^*(x)$ [depicted as a solid line], is not directly sampled, but instead a simpler distribution, $Q^*(x)$ [depicted as a dashed line], such as a normal distribution, is sampled. In regions where $Q^*(x) > P^*(x)$ the samples will *over-represent* $P^*(x)$, and vice versa in regions where $Q^*(x) < P^*$. As a result the relative *importance* of each sample needs to be taken into account, by weighting each sample.

versa when $Q(x)$ is smaller than $P(x)$). To account for this each sample is re-weighted to adjust its importance by the ratio

$$w_r = \frac{P^*(x_r)}{Q^*(x_r)}$$

so then equation 4.19 becomes

$$\hat{\mathbb{E}}(\phi) = \frac{\sum_r w_r \phi(x_r)}{\sum_r w_r}$$

While importance sampling is an improvement over uniform sampling, it will fail to converge in situations where the target distribution contains many separated peaks, and will struggle to explore a high-dimensional space efficiently.

Rejection sampling uses a similar principle to importance sampling, using a *proposal distribution*, $Q(x)$, which can be sampled directly, to generate the samples (see figure 4.8.2 for an illustration of how P and Q relate). The method assumes we know the value of a constant, c such that $cQ^*(x) > P^*(x) \forall x$.

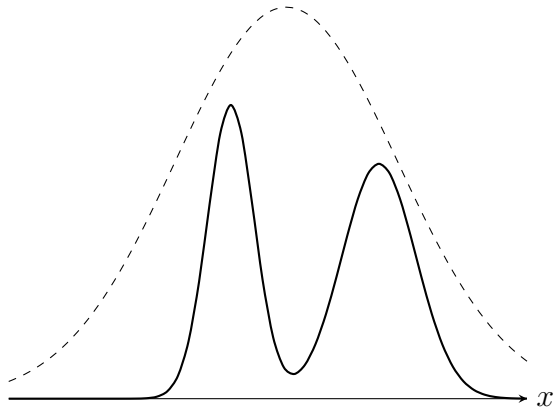


Figure 4.8.2: Similarly to importance sampling, in rejection sampling the arbitrarily complicated distribution, $P^*(x)$ [depicted as a solid line], is not directly sampled, but instead a simpler distribution, the proposal distribution, $Q^*(x)$ [depicted as a dashed line], such as a normal distribution, is sampled. In contrast to importance sampling a constraint is placed on $Q^*(x)$ such that for a constant c $cQ^*(x) > P(x) \forall x$.

This method requires two random numbers to be generated: a sample x is drawn from $Q(x)$, and $cQ(x)$ is calculated. Then a variable u is drawn from the uniform distribution $U(0, cQ^*(x))$. If $u > P^*(x)$ — that is, it lies in the region between $P^*(x)$ and $Q^*(x)$—it is rejected, and discarded. Otherwise, it is accepted, and kept. This method ensures that only points which lie within $P^*(x)$ are retained, preventing over-representation, and also that the density of samples is proportional to $P^*(x)$ thanks to the uniform distribution of samples under $P^*(x)$.

Rejection sampling is fundamentally similar to *Buffon's Needle Problem*, in which needles dropped on floorboards can be used to estimate the value of π, and can be used to evaluate complex integrals outwith probability problems.

Rejection sampling will struggle to converge if the target and proposal distributions are not similar, as the region $[P^*(x), Q^*(x)]$ between the two functions will be large, so the probability of generating samples with $u < P^*(x)$ will be small. The method is also impractical in more than

one-dimension, as similarly, the probability of generating a point within the volume described by $P^*(x)$ will diminish with growing dimensionality.

The deficiencies of these two methods lead to the development of a more sophisticated approach: Markov chain Monte Carlo (MCMC).

4.8.1 Markov-Chain Monte Carlo

As noted previously, rejection sampling struggles to efficiently sample a distribution if the proposal and target distributions are not similar. In order to address this failing, the *Metropolis-Hastings* algorithm constructs a proposal distribution which depends on the sampling location (or more precisely, the current *state* of the sampler). This proposal distribution will often be something simple, like a Normal distribution centred on the current x_t being considered.

As with rejection sampling, a tentative state, x' is drawn from a proposal distribution, $Q^*(x', x_t)$, given the current state, x_t. The ratio

$$a = \frac{P^*(x')}{P*(x_t)} \frac{Q^*(x_t, x')}{Q^*(x', x_t)} \tag{4.20}$$

is evaluated. If $a \geq 1$ the new state is accepted; otherwise the new state is accepted with a probability a. If the new state is accepted it becomes the current state (i.e. $x_{t+1} = x'$); if it is rejected the current state is retained, so $x_{t+1} = x_t$.

In the case that a symmetrical proposal distribution is chosen, such as a normal distribution, the second ratio in equation 4.20 will always be equal to 1, providing a simpler expression for a, which will be consequently faster to evaluate. The behaviour of the Metropolis-Hastings algorithm produces a stochastic process with the Markov property.

In order to improve the computational efficiency of an MCMC algorithm the gradient information of the problem can be taken into account,

which will guide the process to the regions of high probability. These methods, known as *Hamiltonian* MCMC methods can allow faster convergence, and therefore reduce the number of computations required to perform Bayesian inference. The No-U-Turns sampler [219] is an example of such a method which includes various algorithmic refinements to allow the sampler to work efficiently in hierarchical models (see section 4.9) without requiring manual tuning.

4.9 | Hierarchical modelling

Structured probability distributions, as introduced in section 4.5 have the useful property that the posterior distribution can be constructed as the product of a set of independent probability distributions. This structure is frequently useful when describing physical systems, where, for example, we wish to infer the properties of an underlying physical system from a set of individual observations.

An example of such a hierarchical model, used to determine the mean jet opening angle (beaming angle) of sGRBs is presented in chapter 5 and in Williams *et al.* [1], in which a hierarchical approach is taken to determining the probability distribution of the beaming angle via the rates at which observations of sGRBs and BNS events are observed. These are themselves determined from observed quantities, such as the number of observed events, the time over which detections were made, and the false alarm rate of the detection process. A model such as this, which has two layers of inference, is comparatively easy to extend; the inferred beaming angle could, for example, be used as part of the inference of the generating phenomenon.

Hierarchical models are gaining popularity in other areas of GW research, principally black hole population inference [220, 221].

5 Hierarchical Modelling of Gamma Ray Bursts

This chapter is composed principally from the method presented in [1], however the introductory material (in section 5.1) is new compared to that work. Section 5.2 discusses how the beaming angle of short gamma ray bursts may be inferred from the observed astrophysical rate of these events. Section 5.3 discusses what the plausible number of detections during the advanced LIGO observing runs will be, while section 5.4 provides a validation of the technique used to infer the beaming angle from these rates. Section 5.5 provides the bulk of the results from this investigation; figure 5.5.2 illustrates the posterior which can be placed on the beaming angle after an observing run such as advanced LIGO's O2 run. Table 5.5.1 contains a summary of the limits placed over a number of observing scenarios, under a number of different prior assumptions about the fraction of coalescences which produce a jet. This discussion is extended in section 5.6 with an investigation into the beaming angle inference at much higher sensitivities than are currently possible. Section 5.7 provides a summary of the contents of the chapter.

Portions of this chapter, including a number of the figures, have previously appeared [1] in *The Astrophysical Journal*, 2018, The American

Astronomical Society, and is reproduced by permission of the AAS.

5.1 | Short gamma-ray bursts and compact binary coalescences

Gamma ray bursts (GRBs) are highly energetic astrophysical phenomena which were first observed by the VELA nuclear proliferation monitoring satellites [222], and rapidly corroborated [223], in the early 1970s. Twenty years of observations of the events, which occur at around a daily cadence [224], with instruments such as Burst and Transient Source Experiment (BATSE) on the Compton Gamma-Ray Observatory (CGRO) had shown that the distribution of events was isotropic within the sky, but not homogeneous, muddying the waters when trying to identify the progenitor to these events [225]. Instead attempts to classify observed GRBs are based on their duration [226]. This division, with events with a duration in excess of two seconds designated an long gamma ray burst (lGRB), while those with a duration shorter than two seconds are designated short gamma ray bursts (sGRBs), was motivated by the bimodality in the distribution of GRB durations. This distribution, for the GRBs included in the Fermi catalogue [227, 228, 229] is plotted in figure 5.1.1. The T_{90} measure represents the interval between when 5% of the fluence has been detected, and when 95% of the fluence has been detected. By-eye the distribution appears to have a single peak, with a lengthy tail of short-duration events, but there is generally confidence that the distribution in fact best described by a mixture of two Gaussian distributions [230], which correspond to the short and long-duration categories. This classification also demonstrates a bimodality in the distribution of spectral hardness of the events. sGRB were found to be harder, and lGRB softer.

lGRB have come to be associated with core-collapse supernova (CCSN). The observation of SN 1998bw [231], which was observed as an optical

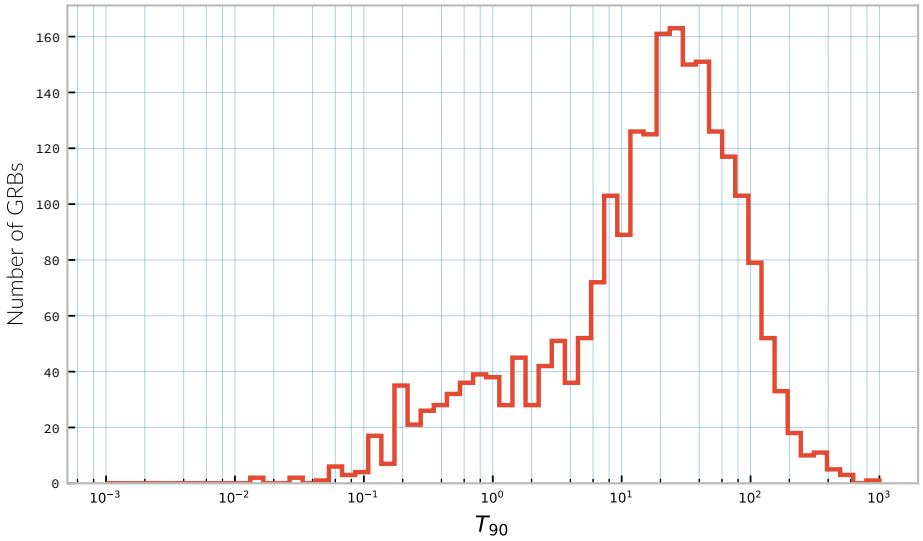


Figure 5.1.1: A histogram of the T_{90} duration of each GRB in the Fermi GRB catalogue [227, 228, 229].

transient within the localisation region of GRB 980425. This transient was shown to be the result of a type Ic supernova. Subsequent numerical modelling [232] of core-collapse supernovae corroborated the connection of the observed supernova and lGRB, although SN 1998bw was unusual, likely through being closer than other GRB, and possibly with a weaker-than-average energy production, allowing the optical transient to be observed. The observation of SN 2003dh, alongside a more normal GRB, and subsequent potential supernova (SN) lGRB associations have added considerably to the evidence of a connection between lGRB and CCSN.

While a connection between SN and lGRB seems plausible, producing the quantity of energy required for an sGRB through this mechanism is impractical in CCSN models. Predictions that an entirely different mechanism may be responsible for these events were made as early as 1984 [233]: namely a binary neutron star (BNS) or neutron star / black hole (NSBH) system. While either of these systems are potentially capable

of producing the quantity of energy required to explain observed sGRB, there is no strong evidence to suggest that all of them will; indeed modelling suggests that a substantial fraction of NSBH systems will be incapable of producing the required energy [234].

The association between sGRB and BNS mergers (or at least, a subset thereof) was cemented by the first observation of a GRB event, GRB 170817A [235], which coincided with a gravitational wave observation: GW170817 [84].

5.1.1 Jet production

At the time of their discovery the properties of sGRBs did not fit any known astrophysical system. The requirement that the events be distributed uniformly in the sky, and have a high observed gamma ray flux emitted over a short period of time made the events difficult to explain. The sharply-peaked nature of the gamma ray lightcurve implied that the progenitors must be small ($< 10^7$ m), moving non-relativistically. A number of early theories included flare stars, antimatter, and neutron star binaries as progenitors (see [236] for a review of these early theories[1]). The favoured theory (especially given the discovery of GRB 170817A, discussed above) for sGRBs is that they are a result of BNS merger events, which are expected to produce the required energy to produce a GRB, provided the emission is beamed.

The determination by BATSE that GRB have a cosmological, rather than Galactic, origin implied that the events must be extremely luminous (and, indeed, the most electromagnetically luminous events observed in the Universe). The fireball model [237], explains the production of energy by the BNS through the kinetic energy of ultra-relativistic particles (or

[1] And also a rather satirical commentary on the state of contemporary astrophysics!

through electromagnetic Poynting flux) being converted to electromagnetic (EM) emission in an optically thin region of material surrounding the merger. This may be a result of interaction with the ISM, or due to interactions within the outflowing material from the merger. This model implies that the majority of emission from the event should be highly beamed, which allows for their extremely high observed luminosity.

5.1.2 Jet geometry

Taking a simple "top-hat" model [2] for the beam profile, which assumes that the emission from an sGRB is concentrated into a conical beam with a half opening angle θ:

$$L(\theta_v) = \begin{cases} L_{\text{uniform}} & \text{if} \quad \theta_v < \theta \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

with θ_v the angle at which the event is viewed, L the luminosity at that angle, and L_{uniform} the luminosity within the beam.

More complex, *structured jet* models exist, such as the Gaussian beam model (introduced for lGRBs in [238]). This model takes the form

$$L(\theta_v) = L_c \exp\left(-\frac{\theta_v^2}{2\theta_c^2}\right) \tag{5.2}$$

where L_c is the luminosity of the jet viewed along its axis, and θ_c is the angle which characterises the width of the beam.

Clearly, if sGRB emission is constrained to a jet, the gamma ray emission will be observed only if the observer is appropriately aligned with the cone of the jet. Since the progenitor of these events is a source of gravitational waves (in the form of a BNS event), which produce near-isotropic emission, it would be reasonable to expect to make detections of BNS

[2] Following the observation of the unusual GRB170817A the efficacy of the top-hat model has been called into question. The top-hat model is, however, easy to work with.

events with gravitational wave (GW) detectors without making coincident gamma ray observations of an associated sGRB. Knowledge of the jet structure allows the construction of a *forward model* which will allow the prediction of the observed rate of sGRB if the jet angle distribution and the rate of BNS events are known.

Understanding how this model might work is easiest by considering a simple game.

1. Make a counter for the number of observed sGRB events, N_{grb}, which is initially set to 0.

2. Now draw N_{bns} values of the viewing angle from a distribution uniform over $\cos(\theta_{\mathrm{v}})$, between $0°$ and $90°$.

3. For each observation draw a value for each of the variables of the luminosity distribution and source distance, D, from appropriate distributions. In the case of the top-hat model this would involve drawing the angle from some distribution limited to the range $0°$ to $90°$ for example.

4. Determine if the observed luminosity, $L(\theta_{\mathrm{v}})/D^2$ is greater than some threshold luminosity, below which the event cannot be observed. If it is, increment the GRB counter, N_{grb}.

Figure 5.1.2 shows the results of playing this game while drawing the beaming angle from various normal distributions truncated between $0°$ and $90°$. This form of the game is particularly simple; we can assume that all events which originate within the volume of space which a GW detector can observe are sufficiently luminous that we will detect their sGRB if viewed along the beam, thus only the jet opening angle affects detectability.

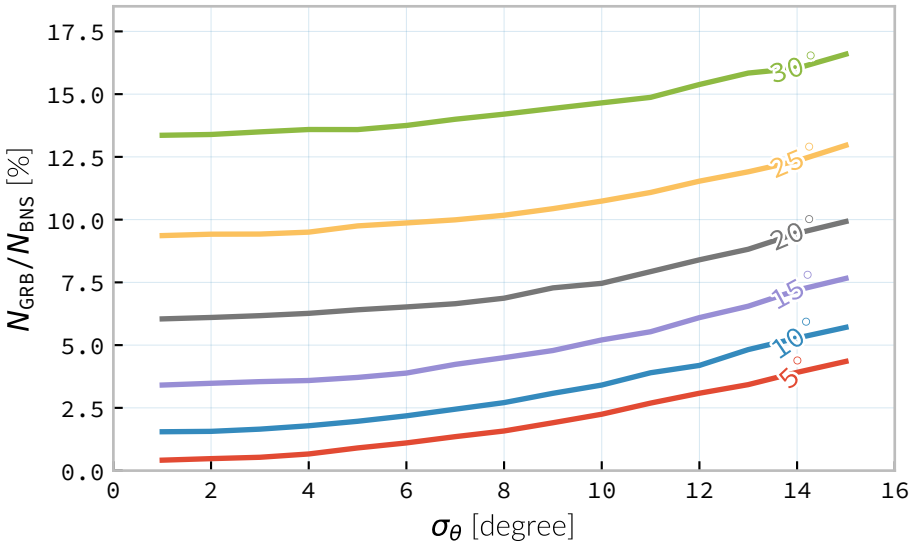


Figure 5.1.2: The expected ratios of observed sGRB to BNS events from a variety of on the beaming angle of a top-hat model. Each distribution is a normal distributions truncated between $0°$ and $90°$; the x axis represents the standard deviation of the distribution, while each line represents a different mean.

Given that it is possible to form a forward model for this scenario, it follows that producing a reverse model, taking advantage of Bayesian inference, should be possible as well.

5.2 | Inferring the beaming angle from astrophysical rates

In this section I will focus on the production of an inferential model for the beaming angle in the top-hat jet model. Provided we assume that all events are sufficiently luminous that we will detect them if we view them along their beam, the simplicity of this model means we can relate the rate at which gamma ray events will be detected, $\mathcal{R}_{\rm grb}$, to the rate at which non-beamed GW events will be detected, $\mathcal{R}$, through the relation

$$\mathcal{R}_{\rm grb} = \epsilon \mathcal{R} \left\langle 1 - \cos\theta \right\rangle , \tag{5.3}$$

where we introduce an efficiency factor, ϵ, to allow for some fraction of merger events to produce no gamma ray emission.

An overview of this approach is as follows:

1. Estimate the posterior probability distribution on the BNS merger rate in the local universe from a number of observed gravitational wave signals and our knowledge of the sensitivity of the detectors. We construct a joint posterior distribution on the BNS rate and the (unknown) probability ϵ that a given merger results in an sGRB.

2. Use equation 5.3, which relates the BNS merger and sGRB rates via the geometry of the beaming angle, to transform the rate posterior probability to a posterior probability on the mean sGRB beaming angle. In this work I consider the observed rate of sGRBs to be constant.

3. Marginalise over ϵ. I choose to consider ϵ a nuisance parameter because, to date, there is no accurate estimate of this parameter and it is not the main focus of our analysis.

In the case that ϵ is very small, and very few BNSs produce a sGRB, then a much larger number of observations will be needed to achieve the same confidence in the measurement of θ than would be required if ϵ was large.

5.2.1 Constructing the BNS rate posterior

In order to make any inference about the sGRB jet angle it is first necessary to determine the rate of BNS events, $\mathcal{R}$. Consequently, an inference step

must be included to determine the probability distribution on this rate, $p(\mathcal{R}|D, I)$, given data, D on GW observations, and other prior information, I.

With the detection of GW170817 in 2017 we now have access to an event upon which to perform inference, however, it is still possible to determine a plausible $\mathcal{R}$ in the absence of detections. This was the scenario during the development of this technique, and I present it here partly to demonstrate its robust nature, and partly to demonstrate how the method may be useful in other multi-messenger scenarios which involve beamed emission. I also present the probability distributions on $\mathcal{R}$ based on the assumptions of observing time and inspiral range presented in the advanced Laser interferometer gravitational-wave observatory (LIGO) *observing scenarios* document [111], which will later be used to determine the future prospects for placing limits on the jet geometry of sGRB events.

This work is not the first attempt to use a comparison of rate to infer information about beam geometry; previously, a comparison of rates was used to place a lower limit on the beaming angle in [239].

GW data analysis search pipelines designed to detect compact binary coalescence (CBC) events, for example `FINDCHIRP` [240], or `PyCBC` [241, 242, 243] identify discrete triggers which are characterised by network signal-to-noise ratio (SNR), ρ_c, which, for the case of BNS searches, indicate the similarity between the detector data and a set of template BNS coalescence waveforms. The measured rate, r, of these events consists of two components: a population of true GW signals, s; and a background rate, b, due to noise fluctuations caused by instrumental and environmental disturbances.

$$r = s + b \begin{cases} s = \text{signal rate} \\ b = \text{background rate.} \end{cases} \tag{5.4}$$

Searches for BNS events are generally conducted as part of an *all-sky* analysis over all of the two-detector coincident data in a given observing run. For these searches, such as those used in [244, 5], the significance of a trigger is determined empirically, by comparing the signal to noise sampled close to its time. A detection requires this significance to be above some predetermined threshold (for example 5σ for GW150914 and GW151226 [105, 107]). I follow the method in [245], which defines a detection as a candidate with an SNR $\rho_c \geq 12$, corresponding approximately to $b = 10^{-2}\,\mathrm{yr}^{-1}$. Since the background rate b is defined, only the signal rate, s, needs to be inferred. In this study I do not consider sub-threshold events (i.e. those with $\rho_c < 12$), and assume that the probability of GW detection from BNS events is not dependent upon the orientation of the source. By not considering sub-threshold events the total volume of space which is observed is effectively reduced, in exchange for maintaining a low background rate of false events, b. In reality there is a greater probability of detecting a face-on BNS event compared to an edge-on BNS event. A face-on BNS event is more likely to have an observable sGRB beam, which may introduce a bias in this method towards broader beam geometries.

By assuming a uniform prior on s and a Poisson process underlying the events, it may be shown (for example in [246]) that the posterior for the signal rate, given a known background rate b and n events observed over a time period T is

$$p(s|n, b, I) = C\frac{T\left[(s+b)T\right]^n e^{-(s+b)T}}{n!}, \tag{5.5}$$

where,

$$C^{-1} = \frac{e^{-bT}}{n!}\int_0^\infty \mathrm{d}(sT)(s+b)^n T^n e^{-sT} \tag{5.6}$$

$$= \sum_{i=0}^{n}\frac{(bT)^i e^{-bT}}{i!}. \tag{5.7}$$

Finally, we can transform the posterior on the *signal* rate to the underlying *coalescence* rate via our knowledge of the sensitivity of the GW analysis. In particular, the signal detection rate is simply the product of the intrinsic coalescence rate $\mathcal{R}$ and the number of BNS mergers which would result in a GW signal with $\rho_c \geq 12$. Expressing the binary coalescence rate in terms of the number of mergers per Milky Way equivalent galaxy (MWEG), per year then we require the number of galaxies N_{G} which may be probed by the GW analysis. At large distances, this is well approximated by [111]:

$$N_G = \frac{4}{3}\pi\left(\frac{\mathcal{D}_{\mathrm{hor}}}{\mathrm{Mpc}}\right)^3 (2.26)^{-3}(0.0116), \tag{5.8}$$

where $\mathcal{D}_{\mathrm{hor}}$ is the horizon distance (defined as the distance at which an optimally-oriented BNS merger yields $\rho_c \geq 12$), the factor of 2.26 results from averaging over sky-locations and orientations, and $1.16 \times 10^{-2}\,\mathrm{Mpc}^{-3}$ is the extrapolated density of MWEG in space.

Finally, the posterior on the binary coalescence rate $\mathcal{R}$ is obtained from a trivial transformation of the posterior on the signal rate s,

$$p(\mathcal{R}|n,T,b,\mathcal{D}_{\mathrm{hor}}) = p(s|n,T,b)\left|\frac{\mathrm{d}s}{\mathrm{d}\mathcal{R}}\right| \tag{5.9}$$

$$= N_G(\mathcal{D}_{\mathrm{hor}})p(s|n,T,b). \tag{5.10}$$

We see that in this approach, the rate posterior depends only on the number of signal detections n, the observation time T, the background rate b, and the horizon distance of the search $\mathcal{D}_{\mathrm{hor}}$. It is precisely these quantities that comprise the detection scenarios outlined in [245]. Before

constructing expected rate posteriors, we outline the transformation from rate to beaming angle.

5.2.2 The sGRB rate

In this work I do not place a prior distribution on the sGRB rate, but assume a fiducial rate, $\mathcal{R}_{\rm grb} = 10\,{\rm Gpc}^{-3}{\rm yr}^{-1}$. A more extensive investigation could attempt to account for the uncertainty in the sGRB rate by placing an astrophysically motivated prior distribution over this quantity.

5.2.3 Constructing the beaming angle posterior

Inferences of the sGRB beaming angle are made from the posterior probability density on the beaming angle $p(\theta|D,I)$ where, as usual, D indicates some set of observations and I unenumerated prior knowledge. Our goal is to transform the measured posterior probability density on the rate $\mathcal{R}$ to a posterior on the beaming angle.

It is possible to transform the joint distribution $p(\theta,\epsilon|D,I)$ using a Jacobian transformation of the joint distribution $p(\mathcal{R},\epsilon|D,I)$:

$$p(\theta,\epsilon) = p(\mathcal{R},\epsilon)\left|\left|\frac{\partial(\mathcal{R},\epsilon)}{\partial(\theta,\epsilon)}\right|\right|, \tag{5.11}$$

(NB, for notational simplicity I will omit the I term herein).

The Jacobian determinant can be computed from equation 5.3. It is then straightforward to marginalize over the efficiency term, ϵ, in order to yield the posterior on θ itself:

$$\begin{aligned} p(\theta) &= \int_\epsilon p(\theta,\epsilon)\,{\rm d}\epsilon & (5.12)\\ &= \int_\epsilon p(\mathcal{R},\epsilon)\left|\left|\frac{\partial(\mathcal{R},\epsilon)}{\partial(\theta,\epsilon)}\right|\right|{\rm d}\epsilon & (5.13)\\ &= \frac{2\mathcal{R}_{\rm grb}\sin\theta\,p(\mathcal{R})}{(\cos\theta-1)^2}\int_\epsilon \frac{p(\epsilon)}{\epsilon}\,{\rm d}\epsilon, & (5.14)\end{aligned}$$

assuming that ϵ and $\mathcal{R}$ are logically independent such that,

$$p(\epsilon, \mathcal{R}) = p(\epsilon|\mathcal{R})p(\mathcal{R}) = p(\epsilon)p(\mathcal{R}). \tag{5.15}$$

It is important to note that the entire procedure of deriving the jet angle posterior is completely independent of the approach used to derive the rate posterior. In the preceding section we adopted a straightforward Bayesian analysis of a Poisson rate which is amenable to a simple application of plausible future detection scenarios; there is no inherent requirement to use that method to derive the rate posterior.

Given the posterior on the rate, $p(\mathcal{R})$, the final ingredient in this approach is the specification of some prior distribution for ϵ. Given the lack of information on the value and distribution of ϵ, three plausible priors were selected, and the distributions on the jet opening angle were inferred under each assumed prior.

The three priors considered are

Delta-function $p(\epsilon) = \delta(\epsilon = 0.5)$; which represents the probability that BNS mergers yield sGRBs is known to be 50% exactly.

Uniform $p(\epsilon) = U(0, 1)$; representing the probability that BNS mergers yield sGRBs may lie anywhere $\epsilon \in (0, 1]$ with equal support in that range.

Jeffreys $p(\epsilon) = \beta(\frac{1}{2}, \frac{1}{2})$; treating the outcome of a BNS merger as a Bernoulli trial in which an sGRB constitutes 'success' and ϵ is the probability of that success, the least informative prior (see 4.3.1). For the Bernoulli distribution, this (Jeffreys) prior is a β-distribution with shape parameters $\alpha = \beta = \frac{1}{2}$.

5.3 | Prospects for beaming angle constraints with advanced LIGO

In this section I will demonstrate the ability of this technique to provide constraints on the beaming angle under a number of plausible observing scenarios for the network of advanced GW detectors. These observing scenarios are derived from the scenarios outlined in [111] which correspond approximately to both the first two observing runs, and planned future observing runs of the network. . An observing scenario essentially consists of an epoch of advanced LIGO operation, which defines an expected search sensitivity (that is, the BNS horizon distance, $\mathcal{D}_{\rm hor}$) and the total observation time T; as well as an assumption on the rate of BNS coalescence in the local universe $\mathcal{R}$. Each observing scenario ultimately results in an expectation for the number of observed GWs from BNS coalescences. For this study, this 'realistic rate' for $\mathcal{R}$ was taken from the method described in [247].

5.3.1 Determining the expected number of observations

Given the observation time and horizon distance of the observation epoch we first compute the 4-volume accessible to the analysis,

$$V_{\rm search} = \frac{4}{3}\pi \left(\frac{\mathcal{D}_{\rm hor}}{2.26}\right)^3 \times \gamma T, \tag{5.16}$$

where the factor 2.26 arises from averaging over source sky location and orientation, T is the observation time and γ is the *duty cycle* for the science run. Following [111], we take $\gamma = 0.5$. For comparison, during the first observing run of advanced LIGO, the two interferometers observed in coincidence achieving a duty cycle $\gamma_{\rm coinc} = 0.41$. Where there is a range in the horizon distances quoted in [111] to account for uncertainty in the

Epoch	T [yr]	$\mathcal{D}_{\rm insp}$ [Mpc]	$V_{\rm search}$ [$\times 10^6 \mathrm{Mpc\,yr^{-1}}$]	Est. BNS Detections
2015–2016	0.25	40–80	0.05–0.4	0.0005–4
2016–2017	0.5	80–120	0.6–2.0	0.006-20
2018–2019	0.75	120–170	3–10	0.04–100
2020+	1	200	20	0.2–200
2024+	1	200	40	0.4–400

Table 5.3.1: Advanced detector era observing scenarios considered in this work. T is the expected duration of the science run and $\mathcal{D}_{\rm insp}$ is the BNS inspiral distance for the sensitivity expected to be achieved at the given epoch, which is equal to $\mathcal{D}_{\rm hor}/2.26$. $V_{\rm search}$ is the sensitive volume of the search, defined by equation 5.16; the final column contains the estimated range of the number of GW detections. Note that the quoted search volume accounts for a network duty cycle of $\sim 80\%$ per detector. These scenarios are derived from those detailed in [111]. While the 2020+ and 2024+ scenarios appear identical in terms of the sensitivity of the detectors, the 2024+ scenario includes a third advanced LIGO detector in India. This expansion of the network is expected to lead to an increase in the network duty cycle, and a corresponding increase in the area of the sky which the network is sensitive to, resulting in a greater volume being searched per year.

sensitivity of the early configuration of the detectors, the arithmetic mean of the lower and upper bounds is used when computing the search volume. Table 5.3.1 lists the details of each observing scenario. The 2015-2016 and 2016-2017 scenarios correspond approximately to the first two advanced LIGO observing runs. The 2018-2019 scenario corresponds to the third observing run, however, since the work in this chapter was prepared, O3 has been extended to a total run-time of 12 months. The 2020+ scenario corresponds to a year of observation with both of the advanced LIGO detectors and Virgo at design sensitivity, with the 2024+ scenario extending this to include a third advanced LIGO detector in India. The increase in the size of the network will lead to an increase in the network duty cycle, and a corresponding increase in the total search volume per year.

Scenario	n	Lower [yr^{-1}]	MAP [yr^{-1}]	Median [yr^{-1}]	Upper [yr^{-1}]
2015–2016	0	0.00	0.45	2.80	11.98
2016–2017	1	0.17	4.07	6.74	19.13
2017–2018	3	1.37	5.88	6.99	15.26
2020+	10	7.30	14.47	15.25	25.25
2024+	20	12.42	20.35	20.65	30.09

Table 5.3.2: Summary of the BNS rate posteriors for each of the observing scenarios which are considered in this work; these posteriors are plotted in figure 5.3.1. Here n is the number of GW events which were assumed to be observed in each scenario, chosen from the ranges in table 5.3.1.

5.3.2 Posterior Results

Having developed a framework in which to infer first the expected BNS rate, and from that the distribution of the jet opening angle, it makes sense to consider how the method is likely to perform as the sensitivity and observing time of the advanced LIGO detectors improves.

Figure 5.3.1 shows the BNS rate posteriors resulting from the observations in the scenarios in table 5.3.1 generated using the procedure described in section 5.2. A number of scenarios have a range of potential inspiral distances, and in each case the median value is used in the analysis, so for the 2015–2016 scenario $\mathcal{D}_{\mathrm{insp}}$ is taken to be $60\,\mathrm{Mpc}$, for example. Likewise an illustrative value of n, the number of expected GW detections, is selected from each range; these are listed in table 5.3.2.

These posteriors, together with the prior distributions described in section 5.2.1 and the observed rate of sGRBs (as described in section 5.1 the rate $\mathcal{R}_{\mathrm{grb}} = 10\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ [248, 249]) is used to derive the corresponding beaming angle posteriors.

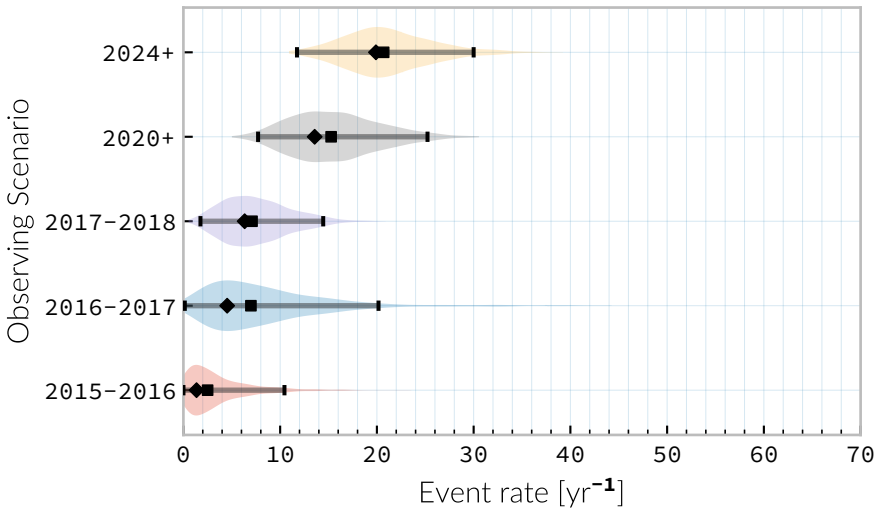


Figure 5.3.1: Posterior probability distribution for the rate of BNS coalescence assuming the scenarios in table 5.3.1. The 95% credible interval is represented with a horizontal line through the centre of the plot, with vertical lines delineating the lower and upper limits; the median is represented by a square marker, and the MAP value is denoted by a diamond. A summary of these values is given in table 5.3.2.

5.4 | Validation

This method is validated by first selecting values of the beaming angle, the sGRB efficiency, and the rate of BNS coalescence. Choosing $\theta = 10°$, $\epsilon = 1$, and the 'realistic' BNS rate $\mathcal{R} = 10^{-6}\,\mathrm{Mpc}^{-3}\mathrm{yr}^{-1}$, the value of the sGRB rate that would correspond to these parameter choices is computed. This *artificial* value for $\mathcal{R}_{\mathrm{grb}}$ is used in equation 5.12 when computing the posterior on the beaming angle, with the understanding that the resulting posterior should yield an inference consistent with the 'true' value $\theta = 10°$.

Figures 5.4.1 and 5.4.2 show the beaming angle posteriors which result from this analysis for the 2015–2016 and 2016–2017 scenarios respectively for each choice of prior distribution on the efficiency parameter. Unsurprisingly, the most accurate constraints arise with the tightest possible

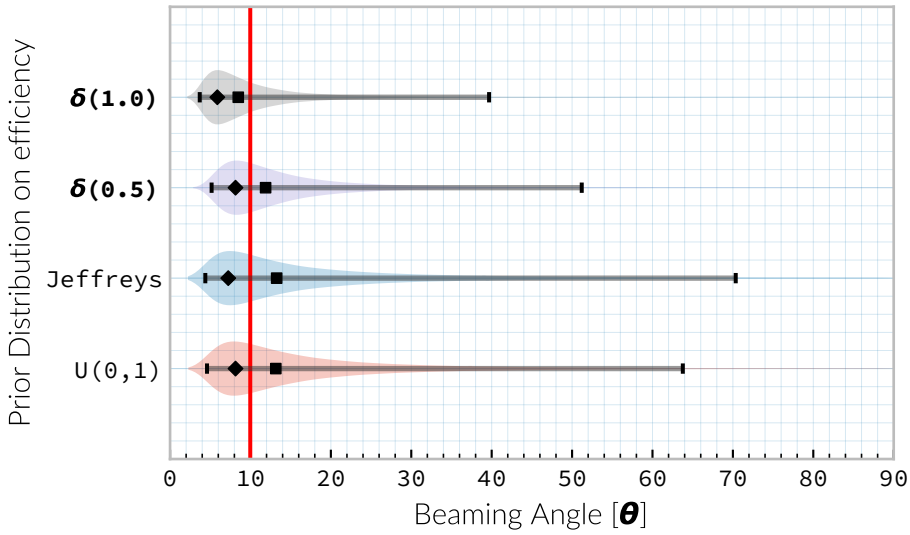


Figure 5.4.1: The posterior probability distributions resulting from the validation analysis described in section 5.4, using the observing time and horizon distance for the 2015–2016 observing scenario (see table 5.3.1). In order to validate the algorithm an artificial scenario was constructed with a known beaming angle by artificially setting an observed sGRB event rate of $36.7\,\mathrm{Gpc^{-3}yr^{-1}}$ to induce a beaming angle of $\theta \approx 10°$.

Prior	Lower [°]	MAP [°]	Median [°]	Upper [°]
$\delta(1.0)$	3.68	5.88	8.45	39.44
$\delta(0.5)$	5.24	8.59	11.89	50.51
Jeffreys	4.38	7.69	13.23	69.74
U(0,1)	4.62	8.14	13.23	63.81

Table 5.4.1: Summary of the beaming angle posteriors from figure 5.4.1, for the 2015–2016 observing scenario, with an artificial sGRB rate imposed to produce a target beaming angle of $\theta = 10°$.

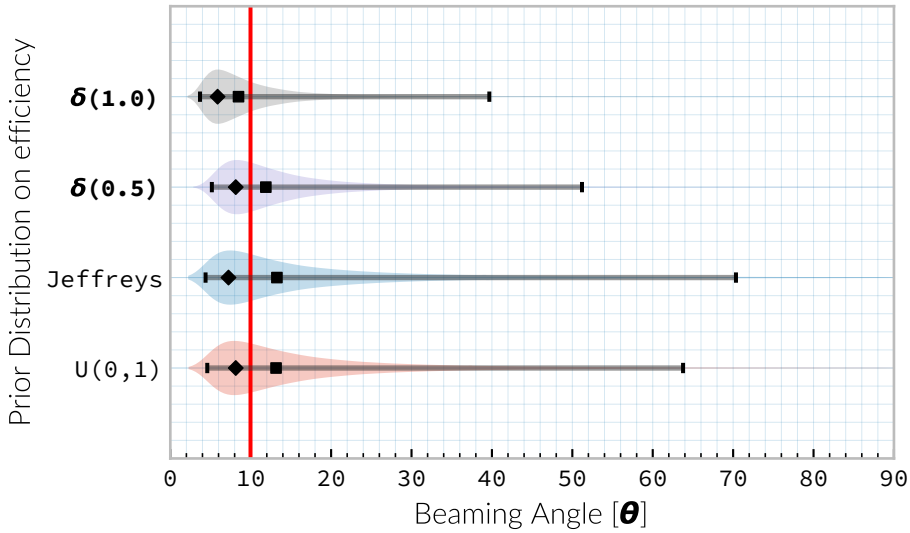


Figure 5.4.2: The posterior probability distributions resulting from the validation analysis described in section 5.4, using the observing time and horizon distance for the 2016–2017 observing scenario (see table 5.3.1). The procedure used to produce figure 5.4.1 was repeated for the observing time and the horizon distance of the 2016–2017 observing scenario, with an observed sGRB event rate of $28.0\,\mathrm{Gpc^{-3}yr^{-1}}$ used to induce a beaming angle of $\theta \approx 10°$. The observed sGRB event rate in this scenario is lower than that used for the 2015–2016 scenario in order to induce the same opening angle despite the greater sensitivity and BNS event rate of this scenario.

Prior	Lower [°]	MAP [°]	Median [°]	Upper [°]
$\delta(1.0)$	4.15	6.78	7.62	21.17
$\delta(0.5)$	6.11	9.50	10.88	27.88
Jeffreys	5.05	9.05	12.21	62.72
U(0,1)	5.12	9.05	11.29	51.04

Table 5.4.2: Summary of the beaming angle posteriors from figure 5.4.2, for the 2016–2017 observing scenario, with an artificial sGRB rate imposed to produce a target beaming angle of $\theta \approx 10°$.

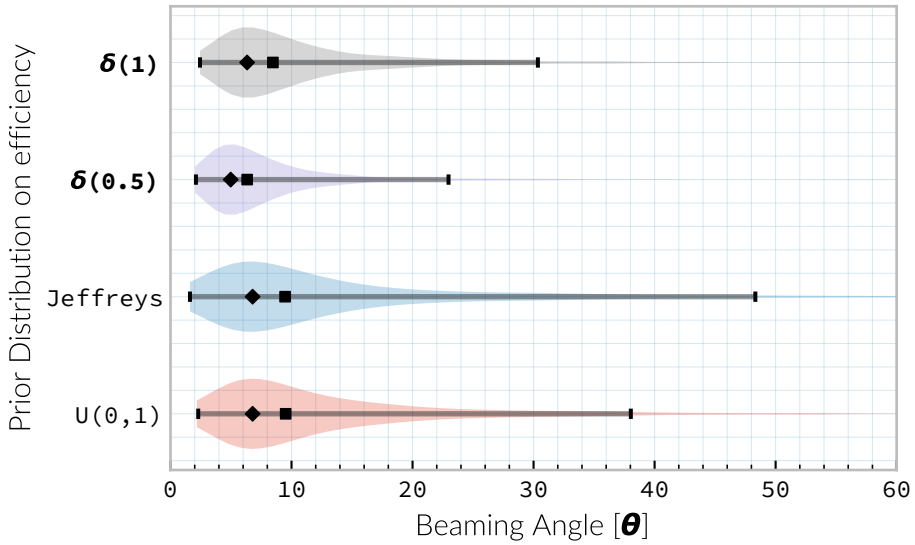


Figure 5.5.1: Beaming angle posteriors using different priors on sGRB efficiency ϵ in the 2015–2016 observing scenario.

constraints on the sGRB efficiency, ϵ. That is, the beaming angle posterior arising from the δ-function prior on ϵ is the narrowest, yielding the shortest possible credible interval. It is worth remembering, however, that an incorrect value of ϵ when using the δ-function prior, would result in a significantly biased posterior, and the inference of the beaming angle would be incorrect. This highlights the necessity of building a suitable representation of ignorance into the analysis.

The similarity of the posteriors which result from the uniform and Jeffreys priors is worth noting, demonstrating that the choice between the least-informative and the indifferent priors leads to only a small difference in the posterior distributions.

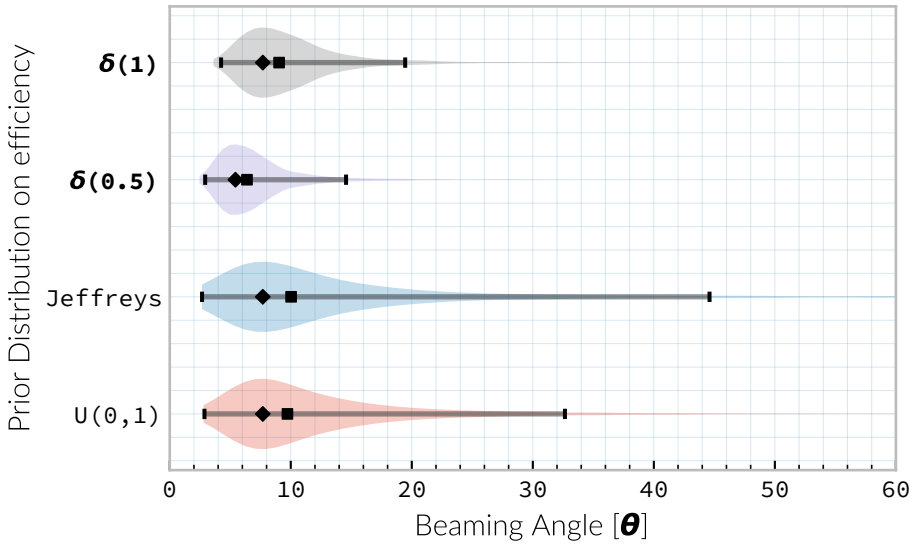


Figure 5.5.2: Beaming angle posteriors using different priors on sGRB efficiency ϵ in the 2016–2017 observing scenario.

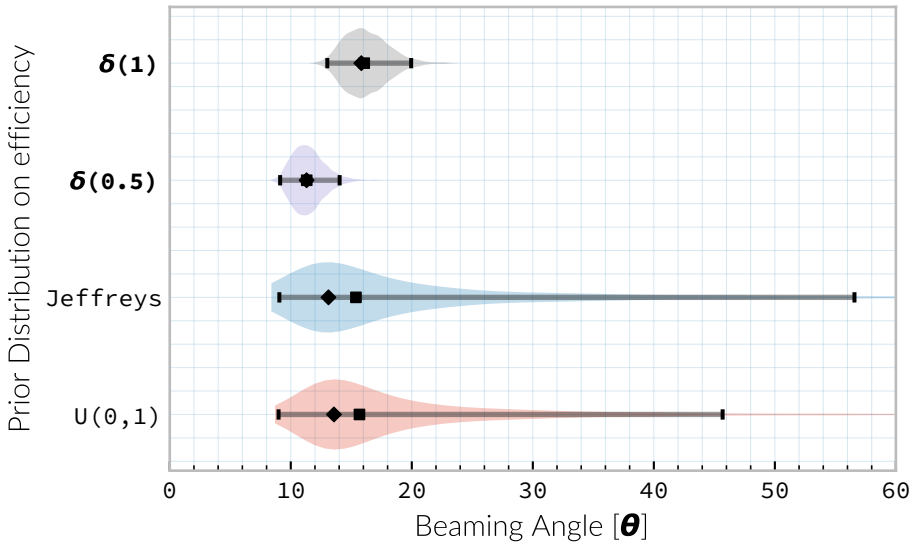


Figure 5.5.3: Beaming angle posteriors using different priors on sGRB efficiency ϵ in the 2024+ observing scenario.

Scenario	Prior	Lower [°]	MAP [°]	Median [°]	Upper [°]
2015–2016	U(0,1)	2.00	5.43	9.24	40.17
	Jeffreys	1.90	5.43	9.50	49.71
	$\delta(1)$	1.76	4.07	5.83	21.04
	$\delta(0.5)$	2.51	5.88	8.22	28.35
2016–2017	U(0,1)	3.09	6.78	9.91	34.23
	Jeffreys	2.85	6.78	9.91	46.93
	$\delta(1)$	2.88	5.43	6.40	14.15
	$\delta(0.5)$	4.06	7.69	9.07	20.05
2018–2019	U(0,1)	6.64	12.66	16.36	46.96
	Jeffreys	6.31	11.76	15.88	57.48
	$\delta(1)$	6.36	9.95	10.97	18.35
	$\delta(0.5)$	8.98	14.02	15.55	26.15
2020+	U(0,1)	8.20	12.66	16.04	44.73
	Jeffreys	7.82	12.21	15.35	56.99
	$\delta(1)$	8.10	10.85	11.12	14.95
	$\delta(0.5)$	11.47	14.92	15.75	21.17
2024+	U(0,1)	9.05	13.12	16.07	45.10
	Jeffreys	8.58	12.21	15.28	56.30
	$\delta(1)$	9.09	11.31	11.30	14.02
	$\delta(0.5)$	12.82	15.83	16.00	19.82

Table 5.5.1: Summary of the beaming angle inferences for each prior in each of the observing scenarios detailed in table 5.3.1. The lower and upper values correspond to the lower and upper bounds of the 95% Bayesian credible interval for each scenario.

5.5 | Results for the advanced LIGO observing scenarios

The posterior distributions on the beaming angle for the first two observing scenarios from table 5.3.1 are plotted as violin plots in figures 5.5.1 and 5.5.2. These observing scenarios are described in table 5.3.1, with the inferred BNS rates for each scenario detailed in table 5.3.2. A fiducial sGRB rate of $\mathcal{R}_{\rm grb} = 10\,{\rm Gpc}^{-3}{\rm yr}^{-1}$ was used for each scenario. These show the

beaming angle posteriors obtained with the various prior distributions listed in section 5.2.3 [3, 4].

Since it is a common assumption in related literature, a prior on the sGRB efficiency which dictates that all BNS produce an sGRB, $p(\epsilon|I) = \delta(\epsilon = 1)$, is also considered in addition to the previous strong δ-function prior.

The 2015-2016 scenario, which corresponds to a three-month observing period in which no BNS signals were detected, provides the least information of the scenarios under consideration, with none of the efficiency priors producing a clear result (the posterior distribution for each of the four efficiency situations is broad). In the 2016-2017 scenario the inference of the beaming angle are also somewhat weak, due to the singular GW detection, and small VT the uncertainties are large enough that the results from each prior are broadly consistent. Both of the posteriors of each of these scenarios are plotted in figures 5.5.1 and 5.5.2 respectively.

In the 2024+ scenario, where the posterior is more peaked, it is clear that the strong δ-function priors lead to inconsistent inferences on the sGRB beaming angle. This can be seen in the plots of each posterior distribution in figure 5.5.3. The much weaker uniform and β distributions, by contrast, are again largely consistent with each other yielding more conservative and robust results, as well as being a more representative expression of our state of knowledge. The inferences drawn from each scenario and each prior are summarised in terms of the MAP measurement and the 95% credible interval around the maximum in table 5.5.1.

One noteworthy feature of these results is the apparent discontinuity in

[3] A note on implementation: rather than directly evaluating the beaming angle posterior in equation 5.12 we choose to sample points from the posterior using an Markov chain Monte Carlo (MCMC) algorithm, implemented using the python package `PyMC3` [250].

[4] While we present the entire posterior for only these two observing scenarios in this section, we provide an overview of all of the observing scenarios in section 5.6.

the inferred beaming angle between the 2016–2017 scenario, and the 2018–2019 scenario. Consulting table 5.3.2 we can see that the median sGRB rate which is inferred for both scenarios is similar, despite the considerable increase in VT between the two scenarios. While this could be taken to imply that the estimate of $n = 3$ BNS events is an underestimate (this corresponds to around half the rate of events that $n = 1$ implies for the 2016–2017 observing scenario), it might equivalently be taken to imply that the observation of one event during the 2016–2017 was simply fortunate. As a result the 2016-2017 observing scenario implies a smaller opening angle to correspond to the larger ratio of BNS rate to sGRB rate. Similarly, the 2015-2016 scenario, in which no BNS events are detected, implies a comparable rate of observed BNS per unit VT to the 2016-2017 scenario, leading to a broadly comparable estimate of the median opening angle in both scenarios.

It is clear from the results presented in table 5.5.1 that under the common assumption that all BNS events should launch a GRB jet that this method allows the most restrictive limits to be placed on the beaming angle; the lower limit placed on the beaming angle from this assumption is comparable in the most sensitive scenario (2024+) for the uniform and Jeffreys priors, however both of these priors produce posterior distributions on the beaming angle which has a long tail, and consequently large upper limits on the beaming angle.

5.6 | Sensitivity beyond the advanced era

While the advanced detectors, such as advanced LIGO are likely to observe a number of BNS events, and a considerable 4-volume of spacetime, the scenarios in 5.3.1 are limited to anticipated sensitivities and event rates within the next decade.

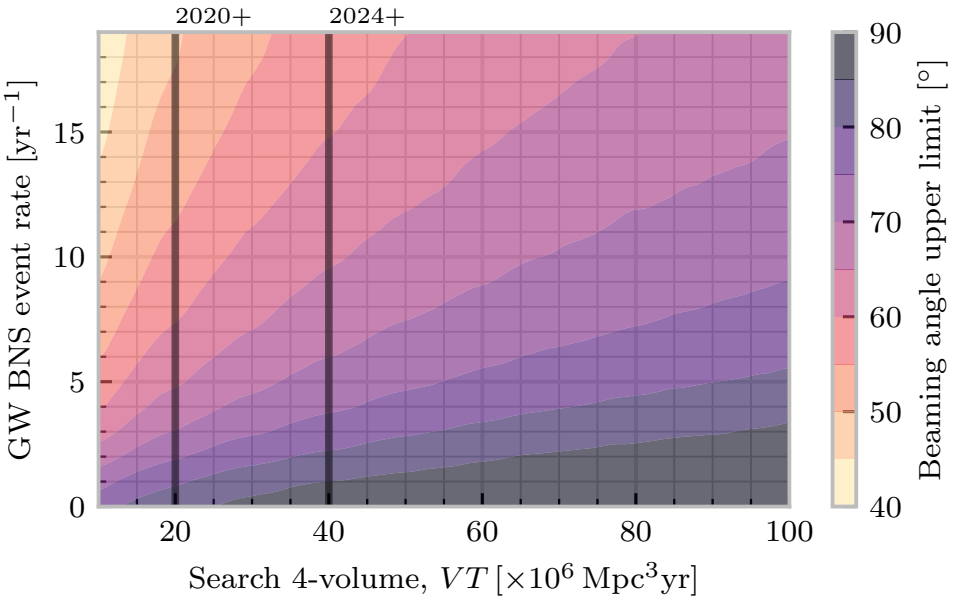


Figure 5.6.1: The upper-bound of the 95% credible interval on the beaming angle as a function of the rate of observed gravitational wave BNS events and the observed search 4-volume, taking a Jeffreys prior on the efficiency of sGRB production from BNS events. The search volumes corresponding to observing scenarios are marked as vertical lines on the plot, with each line assuming that observations are carried out over the period of one year, achieving the search volume outlined in table 5.3.1.

Figures 5.6.1 and 5.6.3 show the upper and lower limits of the 95% confidence region, assuming a Jeffreys prior on the efficiency, as a function of the observed VT and number of BNS events. Similarly, figures 5.6.2 and 5.6.4 show the upper and lower limits of the 95% confidence region assuming all BNS events launch an sGRB. These plots have heavy black lines overlaid to represent the anticipated VT which will be observed by the advanced LIGO network in the 2020+ and 2024+ scenarios.

Similarly to the behaviour seen in the observing scenarios of section 5.5 the upper- and lower-bounds on the beaming angle converge much more rapidly under the assumption that all BNS produce an sGRB compared to when a Jeffreys prior is assumed over the efficiency.

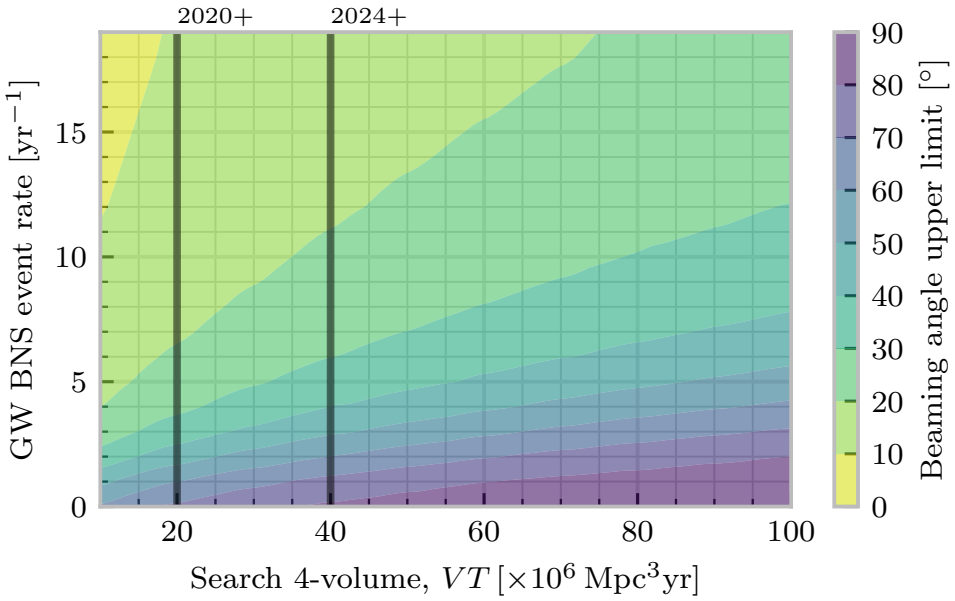


Figure 5.6.2: The upper-bound of the 95% credible interval on the beaming angle as a function of the rate of observed gravitational wave BNS events and the observed search 4-volume, assuming that all BNS events produce an sGRB. The search volumes corresponding to observing scenarios are marked as vertical lines on the plot, with each line assuming that observations are carried out over the period of one year, achieving the search volume outlined in table 5.3.1.

5.7 | Conclusions

The development of this hierarchical Bayesian method for jet angle inference has allowed limits to be placed on the credible region of the sGRB jet beaming angle posterior as a function of the observed number of events and the observed search 4-volume, under a variety of different efficiency conditions. Thanks to the observations of the advanced LIGO detector network during its 2016-2017 observing run, with a single BNS detection, it is possible to place a lower limit of 2.85°, and an upper limit of 46.93° on the jet beaming angle, given an uninformative prior on the efficiency at which BNS events produce observable sGRBs. Assuming that all BNS events produce an observable sGRBs limits narrow to between 2.88° and

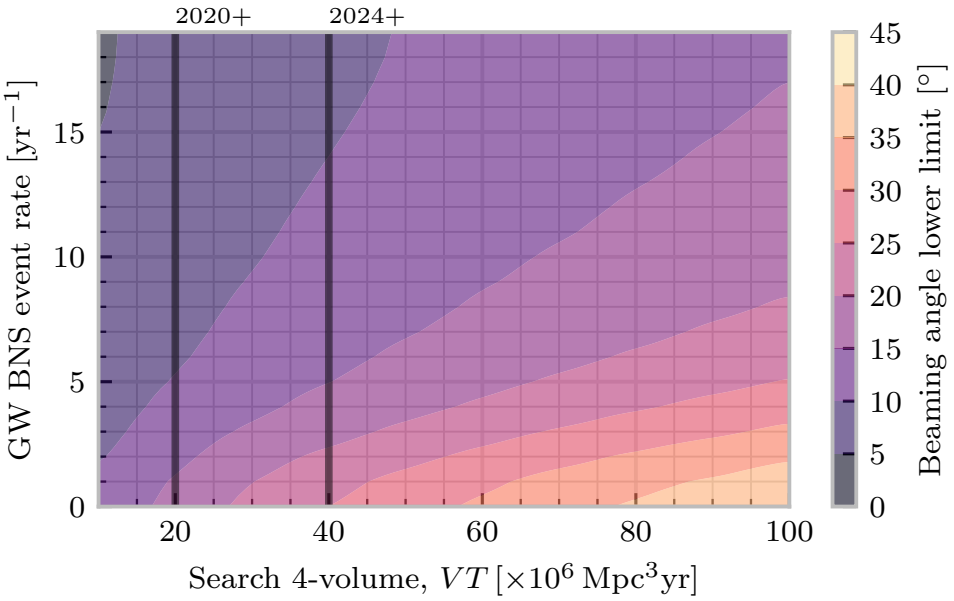


Figure 5.6.3: The lower-bound of the 95% credible limit on the beaming angle as a function of the observed number of events and the observed search 4-volume, taking a Jeffreys prior on the efficiency of sGRB production from BNS events. The search volumes corresponding to observing scenarios are marked as vertical lines on the plot.

14.15°. When advanced LIGO design sensitivity is achieved around 2020 the observation of 10 BNS events in GW is sufficient to place an upper-limit of 56.99° on the jet beaming angle, and can establish the limit on the beaming angle to be between 7.82° and 56.99°, assuming an uninformative prior on the sGRBs production efficiency. These limits narrow to between 8.10° and 14.95° if perfect efficiency is assumed.

In contrasts to previous work, this method incorporates uncertainty in the event rate of BNS signals detected by a GW detector, and also includes uncertainty in the efficiency with which sGRB are produced by these merger events. The limits found from this method are consistent with results using the rates based approach outlined in [239] which finds a lower limit of $2.3°^{+1.7}_{-1.1}$ after analysis of the first LIGO observing run data, under the assumption that all sGRBs are the result of a BNS. Previous methods

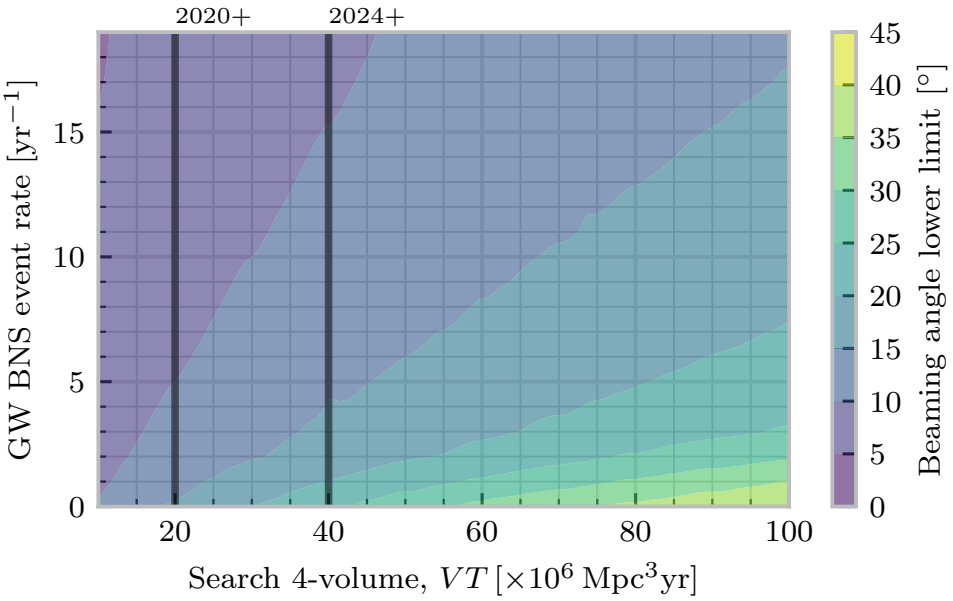


Figure 5.6.4: The lower-bound of the 95% credible limit on the beaming angle as a function of the observed number of events and the observed search 4-volume, assuming that every gravitational wave BNS event produces an sGRB. The search volumes corresponding to observing scenarios are marked as vertical lines on the plot.

have included uncertainty in the BNS rate, but none have performed the analysis using this hierarchical approach. I also believe this is the first work which has attempted to account for the potential efficiency factor, and has presented opening angle estimates using different priors on this quantity.

The work presented in this chapter used a fixed event rate for sGRB in the local universe. This number is, however, uncertain, and it is possible this rate varies outside the immediate vicinity of the Earth. Future work could incorporate this uncertainty by placing an appropriate prior on the sGRB rate, and potentially incorporating considerations based on the sGRB distance into the hierarchical analysis.

The estimates of BNS rate used to demonstrate this method were based on rate posteriors constructed assuming Poisson-distributed events. Given that observational data from GW detectors is now available it is possible

to compute rate posteriors using mock data challenges and real detector noise, allowing for the effects of non-stationarity in the detector noise to be taken into account in the calculation of the rate posterior.

While GRBs are one of the most prominent examples of a beamed emission process in multi-messenger astrophysics, where event rates can be determined through two separate channels, they are by no means unique. This method could be extended easily to situations where beamed particle emission is present, such as high-energy neutrinos, for example. In addition, there are a number of directions this work can be taken in the future. The "top-hat" model has become less favoured since the multimessenger observations associated with GW170817. The sGRB associated with this event was less luminous than would have been expected had the top-hat model been correct [251]. It would therefore be valuable to consider the implication of more complicated "structured" jets on the analysis, and whether it is still possible to make statements using this or a comparable method about the geometrical parameters of those models.

In the analysis presented in this chapter an sGRB rate was chosen and fixed. This was done partly due to the difficulty in determining a suitable prior on this rate at the time the research was conducted, however there has been much work in this area in the last two years [252, 253] and it would be interesting to revisit this assumption, and place a suitable prior distribution on the sGRB rate in order to understand the effect of the uncertainty in this quantity on the inference of the beaming angle.

Additionally, the BNS event rates used in the analysis presented in this chapter are based on anticipated detector sensitivities. The advanced LIGO and Virgo detectors have now completed two observing runs, and the event rate based on the true detector sensitivity and duty cycle can be determined using mock data challenges [5]; this would allow the beaming

[5]See section 3.2 on page 89 for an overview of this technique, in the context of

angle estimate to be based on observational results rather than purely theoretical arguments.

Given the joint observation of GW from a BNS event, and an sGRB during the second observing run of the advanced detectors the prospect of joint GW and EM observations is now a reality. The knowledge that a single event is the source of both GW and BNS provides additional information which the technique presented in this paper is not currently capable of incorporating. Development of the hierarchical analysis to take this into account would likely improve the results of the inference, however this is not likely to be a straight-forward change, since it introduces an additional input datum, the joint-event rate.

The challenges and opportunities which are presented by the arrival of observational multi-messenger results for BNS events make hierarchical analyses such as the one presented in this chapter all the more useful. This modelling technique allows very complex analyses to be built in simpler sections and then connected together, allowing additional effects to be taken into account without a major re-working of the model's implementation in code, for example. The model presented in this chapter is clearly incomplete, but provides one of these sections; a more ambitious project would involve connecting this with analyses of observational GW data and observations from sGRB observatories. Indeed, there is scope to add observations of other EM effects into a much expanded model, including for example, observations in the ultra violet and visible spectra of the kilonova resulting from interactions within the ejecta from the BNS event; and observations across the entire EM spectrum.

---

burst searches. The same principles can be applied to other transient signals, in this case using BNS waveforms rather than burst signals in order to calculate the sensitivity of the detector, taking into account noise non-stationarity and the evolution of the sensitivity through the observing run.

6 Gaussian processes for surrogate modelling

This chapter introduces Gaussian processes, and Gaussian process regression, a probabilistic regression technique which is well suited to producing so-called surrogate models: statistical models which are informed primarily by observed data rather than prior physical knowledge.

This chapter contains mainly introductory and review material, with section 6.7 containing a novel example of a Gaussian process regression model fitting a two-dimensional function. In section 6.1 I introduce the basic ideas of surrogate modelling, and briefly discuss various techniques which exist for producing such models. Then in section 6.2 I introduce the Gaussian process, a probabilistic model, and its use in Gaussian process regression. Section 6.3 discusses the importance of covariance functions for Gaussian process models, and discusses a number of potential choices of covariance function for different situations. A number of examples are provided in section 6.3.1. Section 6.4 discusses the fitting (or training) of Gaussian process regression models. Section 6.5 discusses some of the computational constraints which Gaussian process regression can encounter, and methods which can be used to overcome these. Section 6.6 contains a brief overview of conventional methods used for testing

the accuracy of Gaussian process regression models. Section 6.7 contains a novel example of a Gaussian process regression model being used to determine contour lines on a geographical map, and shows examples of outputs from a range of different covariance functions.

6.1 | Surrogate modelling

Many experimental scenarios in science and engineering are expensive, laborious, or both to perform, and therefore difficult or off-putting to repeat. The ability to perform such an experiment with a specific set of parameters may however be valuable.

While in engineering the difficulties of performing repeatative experiments may relate to the nature of the experiment (which might, for example, be destructive to a tested material), in astrophysics we are more frequently hampered by the computational expense of running a numerical simulation.

A number of techniques exist to perform interpolation between experimentally measured data. If a well-defined physical theory is known to explain the observations an appropriate function can be derived to make predictions outwith the sampled parameter space for the experiment, and this can be calibrated (fitted) to the data.

Surrogate models are often desirable in situations where little physical intuition is available (or where the underlying physics is extremely complicated). These attempt to model the observed data with limited assumptions about the form of the generating function. This approach, of making limited assumptions regarding the structure of the underlying model often results in these techniques being considered a sub-division of the nebulous field of *machine learning*.

Polynomial response surfaces Perhaps the most straight-forward means of producing a surrogate model is to assume that the generating process for measured data can be approximated by a polynomial function. In the case of many complex physical processes there may be a number of coefficients, leading to a polynomial (hyper-)surface. This method, formally introduced in 1951 [254], is often referred to

in the literature as the response surface methodology (RSM), and is suitable for approximating generating models which have smooth variation, and non-oscillatory behaviour.

Support vector machine regression (SVR) extend support vector machine classification to handle regression problems [255]. This approach affords greater flexibility than polynomial model, as kernel methods can be used to perform non-linear fitting, projecting the data into a feature space.

Gaussian process (GP) regression which will be the focus of this chapter, is a technique which possesses similarities to SVR, in that it makes use of kernel functions to project the data into a feature space, however Gaussian process regression (GPR) is a fundamentally Bayesian regression model, and as such provides not only a regressor to data, but also a measure of the fit's uncertainty.

Artificial neural networks Neural network techniques, including the use of convolutional neural networks, and deep learning techniques, are capable of producing non-linear fits to data. See [256] for an example of this technique used as a surrogate model, in this case to evaluate the strength of various concrete mixes.

6.2 | Gaussian Processes

Consider a regression problem with a set of data

$$\mathcal{D} = \{(\boldsymbol{x}_i, y_i), i \in 1, \ldots, n\}$$

which is composed of n pairs of inputs, $\boldsymbol{x}_i$, which are vectors which describe the location of the datum in parameter space, which are the inputs for the problem, and y_i, the outputs. The outputs may be noisy; in

this work I will only consider situations where the noise is additive and Gaussian, so

$$y_i(\boldsymbol{x}_i) = f(\boldsymbol{x}_i) + \epsilon_i, \quad \text{for} \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \tag{6.1}$$

where σ is the standard deviation of the data noise, and f is the (latent) generating function of the data.

This regression problem can be addressed using *Gaussian processes*:

Definition 6.2.1. A Gaussian process (GP) is a collection of random variables, any finite number of which have a joint Gaussian distribution [257].

Where it is more conventional to consider a prior over a set of, for example, real values, such as a normal distribution, the Gaussian process forms a prior over the functions, f, from equation 6.1, which might form the regression fit to any observed data. This assumes that the values of the function f behave as

$$p(\boldsymbol{f}|\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \mathcal{N}(0, \mathsf{K}) \tag{6.2}$$

where K is the covariance matrix of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$, which can be calculated with reference to some *covariance function*, k, such that $K_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. Note that I have assumed that the GP is a *zero-mean* process; this assumption is frequent within the literature. While this prior is initially untrained it still contains information about our preconceptions of the data through the form of the covariance function. For example, whether or not we expect the fit to be smooth, or periodic. Covariance functions will be discussed in greater detail in section 6.3.

By providing training data we can use Bayes theorem to update the Gaussian process, in the same way that the posterior distribution is updated by the addition of new data in a standard Bayesian context, and a posterior on the set of all possible functions to fit the data is produced. Thus, for

a vector of test values of the generating function $\boldsymbol{f}_\star$, the joint posterior $p(\boldsymbol{f}, \boldsymbol{f}_*|\boldsymbol{y})$, given the observed outputs $\boldsymbol{y}$ can be found by updating the GP prior on the training and test function values $p(\boldsymbol{f}, \boldsymbol{f}_*)$ with the likelihood $p(\boldsymbol{y}|\boldsymbol{f})$:

$$p(\boldsymbol{f}, \boldsymbol{f}_*|\boldsymbol{y}) = \frac{p(\boldsymbol{f}, \boldsymbol{f}_*)p(\boldsymbol{y}|\boldsymbol{f})}{p(\boldsymbol{y})}. \tag{6.3}$$

Finally the (latent) training-set function values, $\boldsymbol{f}$ can be marginalised out:

$$p(\boldsymbol{f}_*|\boldsymbol{y}) = \int p(\boldsymbol{f}, \boldsymbol{f}_*|\boldsymbol{y})\,\mathrm{d}\boldsymbol{f} = \frac{1}{p(\boldsymbol{y})}\int p(\boldsymbol{y}|\boldsymbol{f})p(\boldsymbol{f}, \boldsymbol{f}_*)\,\mathrm{d}\boldsymbol{f} \tag{6.4}$$

We can take the mean of this posterior in the place of the "best fit line" which other techniques produce, and then use the variance to produce an estimate of the uncertainty of the prediction.

Both the prior $p(\boldsymbol{f}, \boldsymbol{f}_*)$ and the likelihood $p(\boldsymbol{y}|\boldsymbol{f})$ are Gaussian:

$$p(\boldsymbol{f}, \boldsymbol{f}_*) = \mathcal{N}(\mathbf{0}, \mathsf{K}^+), \quad \text{and} \quad p(\boldsymbol{y}|\boldsymbol{f}) = \mathcal{N}(\boldsymbol{f}, \sigma^2\mathsf{I}) \tag{6.5}$$

with

$$\mathsf{K}^+ = \begin{bmatrix} \mathsf{K}_{\boldsymbol{f},\boldsymbol{f}} & \mathsf{K}_{\boldsymbol{f},\boldsymbol{f}_*} \\ \mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{f}} & \mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{f}_*} \end{bmatrix}, \tag{6.6}$$

and I the identity matrix, and σ is the standard deviation of additive noise in the data.

This leaves the form of the marginalised posterior being analytical:

$$p(\boldsymbol{f}_*|\boldsymbol{y}) = \mathcal{N}(\boldsymbol{\mu}, \Sigma) \tag{6.7}$$

for

$$\boldsymbol{\mu} = \mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{f}}(\mathsf{K}_{\boldsymbol{f},\boldsymbol{f}} + \sigma^2\mathsf{I})^{-1}\boldsymbol{y}, \tag{6.8}$$

and

$$\Sigma = \mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{f}_*} - \mathsf{K}_{\boldsymbol{f},\boldsymbol{f}_*}(\mathsf{K}_{\boldsymbol{f},\boldsymbol{f}} + \sigma^2\mathsf{I})^{-1}\mathsf{K}_{\boldsymbol{f},\boldsymbol{f}_*}). \tag{6.9}$$

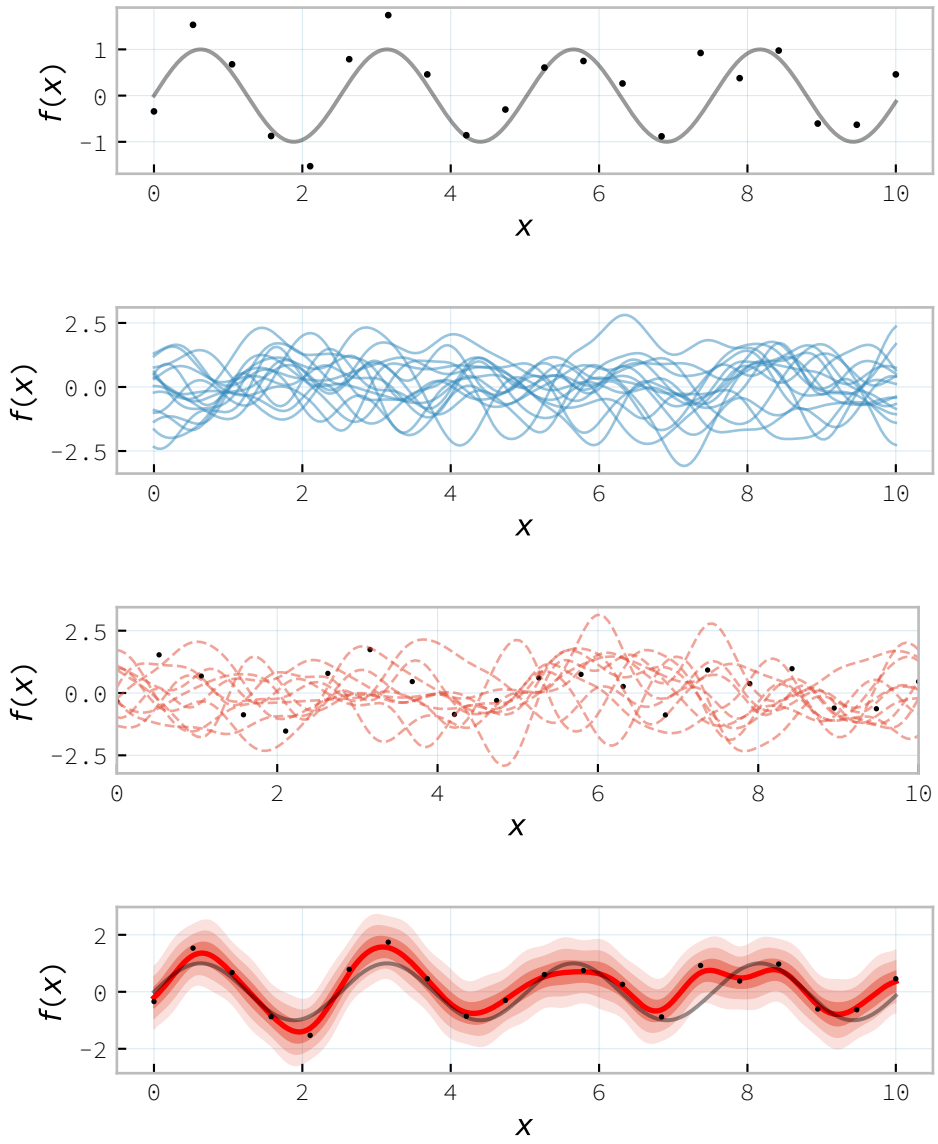


Figure 6.2.1: The conditioning of a Gaussian process, starting with data with additive Guassian noise generated from a sine function (grey line) [top row], and a Gaussian process prior [second row]. Individual draws from the posterior distribution of the Gaussian process are shown in the third row, and the mean draw from the posterior is shown as the heavy red line in the bottom row, with the function which generated the data overlayed in grey, and the 1, 2, and 3-sigma confidence regions plotted as shaded areas around the mean draw.

Figure 6.2.1 shows visually how a one-dimensional regressor can be created using an GP method, starting from a GP prior and (noisy) data. The first step, depicted on the first row, is an example of raw training data (containing additive Gaussian noise) which is suitable for training a Gaussian process. In this example the input data (x-axis) are 1-dimensional, although GPs are also capable of handling multi-dimensional data. Here the generating function is plotted as a grey line. Then we choose a covariance function for the GP, in this case an exponential-quadratic covariance function (covariance functions are discussed in detail in section 6.3). The Gaussian process containing no data forms our prior probability distribution. In the second row of figure 6.2.1 10 draws from the prior distribution are plotted. The process of "training" the GP is discussed in detail in section 6.4. A prior distribution is placed over the σ parameter (see equation 6.5), and the GP is trained to find the most probable value for the σ parameter and the hyperparameters of the covariance function. For this example the prior placed on σ is $\mathcal{N}(0.5, 0.2)$, and the priors on the hyperparameters are flat. The trained Gaussian process can then be sampled multiple times to produce multiple different potential fitting functions. In the third row of figure 6.2.1 10 draws from the GP posterior are displayed. We can also take the mean and the covariance of the GP, and produce a single "best-fit" with confidence intervals, which is depicted in the fourth row of figure 6.2.1. Where, again, the original generating function for the data is shown as a grey line. The mean function produced by the GP manages to reproduce a function which oscillates in a way similar to the generating sine function, however the presence of a considerable amount of noise in the data, which is accounted for through the σ term in the GP, prevents the function from being recovered completely faithfully.

The mean and variance of this posterior distribution can be used to form a regressor for the data, $\mathcal{D}$, with the mean taking the role of a "line-of-

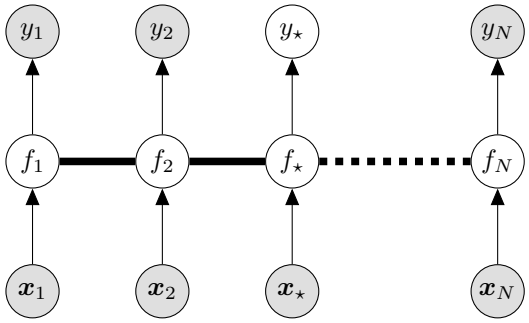


Figure 6.2.2: A graphical model of a Gaussian process, represented as a chain graph. The inputs (on the bottom row) are all observed quantities, while outputs are observed only at the location of training points. The latent variables, f from the Gaussian field (the heavy black line connecting these nodes indicates that they are fully connected) connect the two, and so any given observation is independent of all other nodes given its connected latent f variable. Thus the marginalisation (removal) or addition of input nodes to the GP does not change the distribution of the other variables.

best-fit" in conventional regression techniques, while the variance describes the goodness of that fit.

A graphical model of a GP is shown in figure 6.2.2 which illustrates an important property of the GPs model: the addition (or removal) of any input point to the GP does not change the distribution of the other variables. This property allows outputs to be generated at arbitrary locations throughout the parameter space.

6.3 | Covariance Functions

The covariance function defines the similarity of a pair of data points, according to some relationship with suitable properties. The similarity of input data is assumed to be related to the similarity of the output, and therefore the more similar two inputs are the more likely their outputs are to be similar.

As such, the form of the covariance function represents prior knowl-

edge about the data, and can encode understanding of effects such as periodicity within the data.

Definition 6.3.1 (Stationary covariance function)**.** Given two points, $\boldsymbol{x}$ and $\boldsymbol{x}'$ in a parameter space, a stationary covariance function is a function $f(\boldsymbol{x} - \boldsymbol{x}')$, and which is thus invariant to translations in the input space.

That is, the function depends on the separation of the points, and not their position.

Definition 6.3.2 (Isotropic Covariance Function)**.** Given two points, $\boldsymbol{x}$ and $\boldsymbol{x}'$ in a parameter space, if a covariance function is a function of the form $f(|\boldsymbol{x} - \boldsymbol{x}'|)$ then it is isotropic, and invariant under all rigid motions.

Thus such a covariance function depends only on the separation between the points, and not the direction between them.

A covariance function which is both stationary and isotropic has the property that it can be expressed as a function of a single variable, $r = |\boldsymbol{x} - \boldsymbol{x}'|$ is known as a radial basis function (RBF). Functions of the form $k : (\boldsymbol{x}, \boldsymbol{x}') \to \mathbb{C}$, for two vectors $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ are often known as *kernels*, and I will frequently refer interchangably to covariance functions and kernels where the covariance function has this form.

For a set of points $\{\boldsymbol{x}_i | i = 1, \ldots, n\}$ a kernel, k can be used to construct the gram matrix, $K_{i,j} = k(x_i, x_j)$. If the kernel is also a covariance function then K is known as a *covariance matrix*.

For a kernel to be a valid covariance function for a GP it must produce a positive semidefinite covariance matrix K. Such a matrix, $\mathsf{K} \in \mathbb{R}^{n \times n}$ must satisfy $\boldsymbol{x}^\intercal \mathsf{K} \boldsymbol{x} \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

6.3.1 Example covariance functions

One of the most frequently encountered covariance functions in the literature is the squared exponential (SE) covariance functions [257]. Perhaps as a result of its near-ubiquity this kernel is known under a number of similar, but confusing names (which are often inaccurate). These include the *exponential quadratic*, *quadratic exponential*, *squared exponential*, and even *Gaussian* covariance function.

The reason for this is its form, which closely resembles that of the Gaussian function:

$$k_{\mathrm{SE}}(r) = \exp\left(-\frac{r^2}{2l^2}\right), \tag{6.10}$$

for r the Euclidean distance of a datum from the centre of the parameter space, and l is a scale factor associated with the axis along which the data are defined.

The SE function imposes strong smoothness constraints on the model, as it is infinitely differentiable. This covariance function is therefore well-suited to modelling data which is generated by smooth processes without discontinuities.

The scale factor, l in equation 6.10, also known as its *scale-length* defines the size of the effect within the process. This characteristic length-scale can be understood [258, 257] in terms of the number of times the GP should cross some given level (for example, zero). Indeed, for a GP with a covariance function k which has well-defined first and second derivatives the expected number of times, N_u the process will cross a value u is [257]

$$\mathbb{E}(N_u) = \frac{1}{2\pi}\sqrt{-\frac{k''(0)}{k(0)}}\exp\left(-\frac{u}{2k(0)}\right) \tag{6.11}$$

A zero-mean GP which has an SE covariance structure will then cross zero $1/(2\pi l)$ times on average.

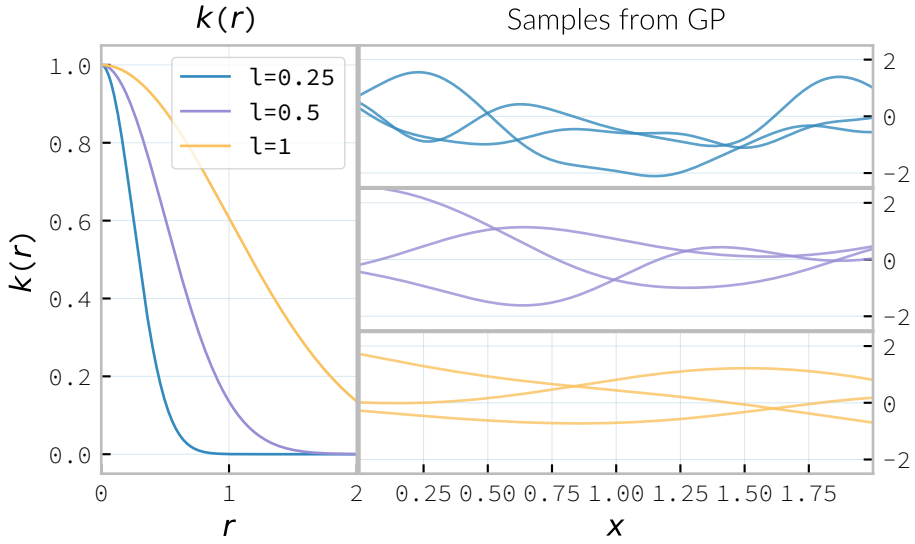


Figure 6.3.1: The **squared exponential** covariance function (defined in equation 6.10). The panel on the left depicts the value of the kernel as a function of $r = (|\boldsymbol{x} - \boldsymbol{x}'|)$, at a number of different length scales ($l = 0.25, 0.5, 1.0$) while the panel on the right contains draws from Gaussian processes using SE covariance with the same length scales as the left panel.

Examples of the SE covariance function, and of draws from a Gaussian process prior which uses this covariance function are plotted in figure 6.3.1 for a variety of different scale lengths.

For data which is not generated by a smooth function a suitable covariance function may be the exponential covariance function, k_{EX}, which is defined

$$k_{\text{EX}} = \exp\left(-\frac{r}{l}\right), \tag{6.12}$$

where r is the pairwise distance between data and l is a length scale, as in equation 6.10.

In contrast to the SE covariance function, the exponential covariance function's value drops-off rapidly near zero (as can be seen in the left panel

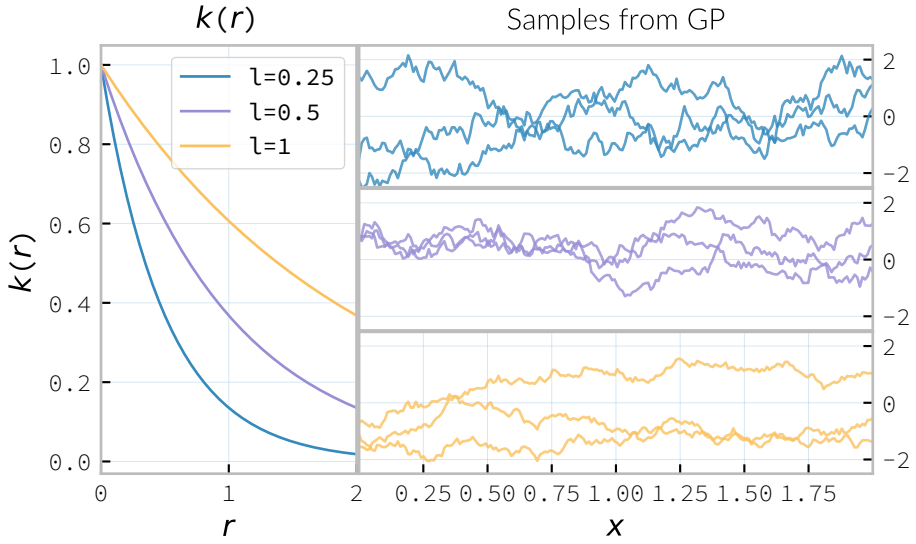


Figure 6.3.2: The **exponential** covariance function (defined in equation 6.12). The panel on the left depicts the value of the kernel as a function of $r = (|\boldsymbol{x} - \boldsymbol{x}'|)$, at a number of different length scales ($l = 0.25, 0.5, 1.0$) while the panels on the right contain draws from Gaussian processes using an exponential covariance with the same length scales as the left panel.

of figure 6.3.2), allowing it to model rapid variation over short scales, making it suited to modelling data generated by non-smooth functions.

Examples of the exponential covariance function, and of draws from a Gaussian process prior which uses this covariance function are plotted in figure 6.3.2 for a variety of different scale lengths. The behaviour of this kernel is strongly affected by the covariance function's rapid drop-off close to zero; compared to the other examples of covariance function in this section.

For data generated by functions which are smooth, but not necessarily infinitely differentiable, as in the case of the SE covariance function, we may turn to the Matérn family of covariance functions, which take the form

$$k_{\text{Mat}}(r) = \frac{1}{2^{\nu-1}\Gamma\nu}\left(\frac{\sqrt{2\nu}}{l}\right)^{\nu} K_{\nu}\left(\frac{\sqrt{2\nu}}{l}r\right), \tag{6.13}$$

for K_ν the modified Bessel function of the second kind, and Γ the gamma function. As with the previous two covariance functions l is a scale length parameter, and r the distance between two data. A GP which has a Matérn covariance function will be $(\lceil x \rceil - 1)$-times differentiable.

While determining an appropriate value of ν during the training of the GP is possible, it is common to select a value *a priori* for this quantity. $\nu = 3/2$ and $\nu = 5/2$ are common choices as K_ν can be determined simply, and the covariance functions are analytic.

The case with $\nu = 3/2$, commonly referred to as a Matérn-$3/2$ kernel, then becomes

$$k_{\text{M32}}(r) = \left(1 + \frac{\sqrt{3}d}{l}\right)\exp\left(-\frac{\sqrt{3}d}{l}\right). \tag{6.14}$$

Examples of this covariance function, and example draws from a GP using it as a covariance function are plotted in figure 6.3.3.

Similarly, the Matérn-$5/2$ is the case where $\nu = 5/2$, taking the form

$$k_{\text{M52}}(r) = \left(1 + \frac{\sqrt{5}d}{l} + \frac{5d^2}{3l^2}\right)\exp\left(-\frac{\sqrt{5}d}{l}\right). \tag{6.15}$$

Again, examples of this covariance function, and example draws from a GP using it as a covariance function are plotted in figure 6.3.4.

Data may also be generated from functions with variation on multiple scales. One approach to modelling such data is to use a GP with **rational quadratic** covariance. This covariance function represents a scale mixture of RBF covariance functions, each with a different characteristic length scale. The rational quadratic covariance function is defined as

$$k_{\text{RQ}}(r) = \left(1 + \frac{r^2}{2\alpha l^2}\right)^{-\alpha}, \tag{6.16}$$

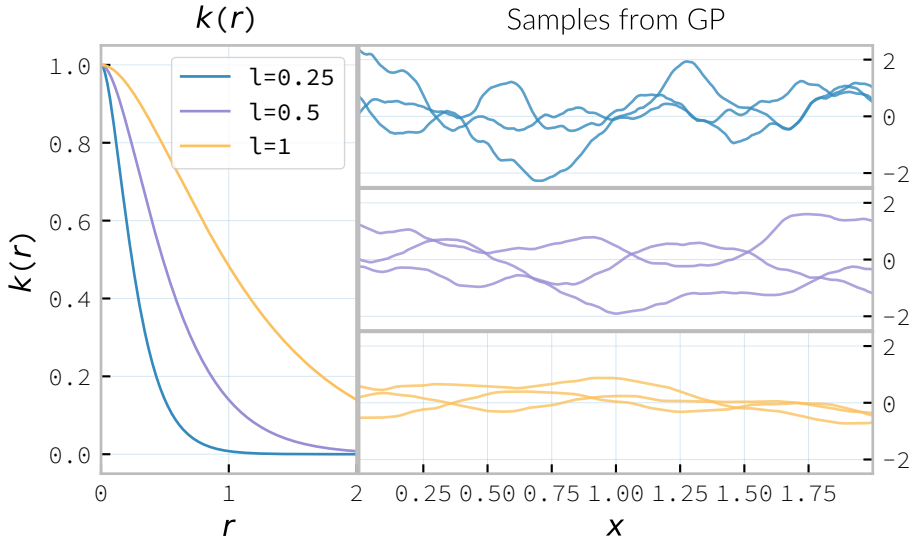


Figure 6.3.3: The **Matérn-**$3/2$ covariance function (defined in equation 6.13, with $\nu = 3/2$). The panel on the left depicts the value of the kernel as a function of $r = (|\boldsymbol{x} - \boldsymbol{x}'|)$, at a number of different length scales ($l = 0.25, 0.5, 1.0$) while the panels on the right contain draws from Gaussian processes using a Matérn-$3/2$ covariance with the same length scales as the left panel.

where α is a parameter which controls the weighting of small-scale compared to large-scale variations, and l and r are the overall length scale of the covariance and the distance between two data respectively. Examples of this function, at a variety of different length scales and α values, and draws from GPs which use these functions are plotted in figure 6.3.5.

This summary of potential covariance functions for use with a GP is far from complete (see [257] for a more detailed list). However, these four can be used or combined to produce highly flexible regression models, as they can be added and multiplied as normal functions.

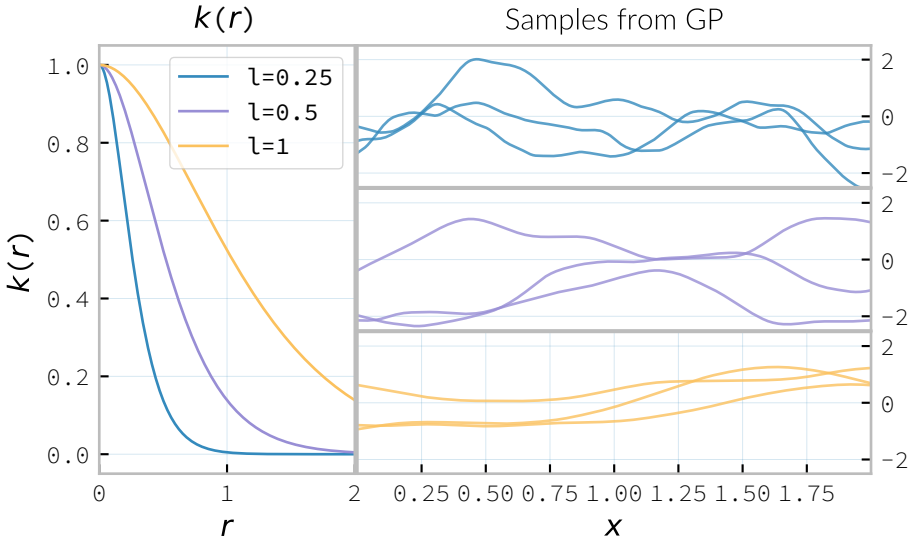


Figure 6.3.4: The **Matérn-**$5/2$ covariance function (defined in equation 6.13, with $\nu = 5/2$). The panel on the left depicts the value of the kernel as a function of $r = (|\boldsymbol{x} - \boldsymbol{x}'|)$, at a number of different length scales ($l = 0.25, 0.5, 1.0$) while the panels on the right contain draws from Gaussian processes using Matérn-$5/2$ covariance functions with the same length scales as the left panel.

6.3.2 Kernel algebra

It is possible to define new kernels from the standard set through a series of defined operations.

Consider two covariance functions, f_1 and f_2, then

Definition 6.3.3 (Kernel Addition)**.** If f_1 and f_2 are both kernels, then $f = f_1 + f_2$ is also a kernel.

Definition 6.3.4 (Kernel Multiplication)**.** If f_1 and f_2 are both kernels, then $f = f_1 \times f_2$ is also a kernel.

We can think of the sum of two kernels as representing the possibility that the data be described by one component kernel or another. As such

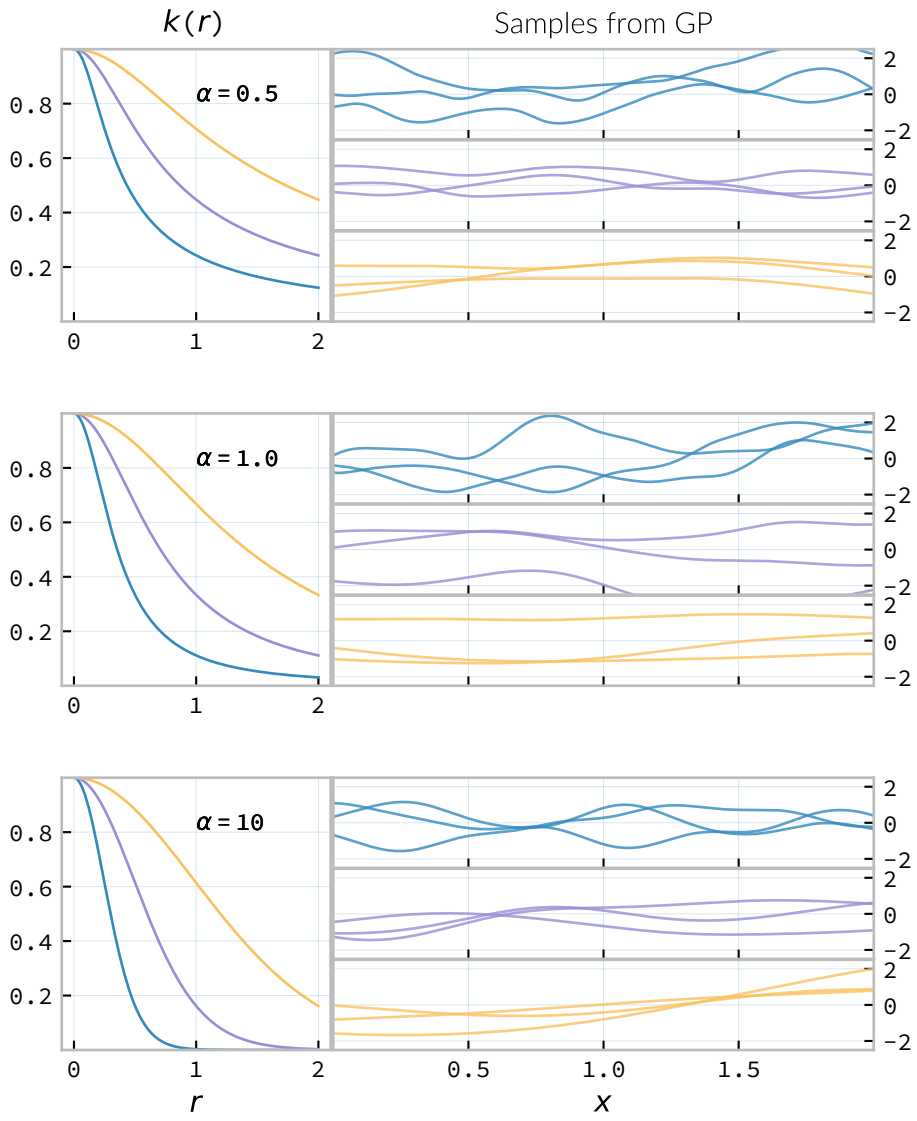


Figure 6.3.5: The **rational quadratic** covariance function (defined in equation 6.16). The panel on the left depicts the value of the kernel as a function of $r = (|\boldsymbol{x} - \boldsymbol{x}'|)$, at a number of different length scales ($l = 0.25, 0.5, 1.0$) while the panel on the right contains draws from Gaussian processes using rational quadratic covariance with the same length scales as the left panel.

addition represents the logical OR operation. Similarly the product of two kernels represents the logical AND operation between the two.

We can use these two operations to form an arbitrarily complicated kernel structure, and to allow inference to be conducted over multiple dimensions. Different kernels can be used to model different aspects of the variation within the input data. For example, the training data may be known to be periodic in one dimension, or to have white noise properties in another. Here I adopt the convention from [259] and omit the hyperparameters from the description of the kernel. I also extend the notation to allow kernels with multiple input dimensions to be described, with superscript indices indicating the dimensions of the training data which the kernel applies to.

As a concrete example, for a kernel function in which the zeroth dimension is described by a SE kernel, but the first, second, and third dimensions are described by a rational quadratic kernel the kernel could be described as

$$k = \mathsf{SE}^{(0)} \times \mathsf{RQ}^{(1,2,3)} \tag{6.17}$$

A list of the symbols for each covariance function is given in table 6.3.1, and definitions of the kernels are given at the end of the chapter.

For example, we may be able to model a yearly growing trend which contains a seasonal variation with a combination of a linear and a periodic kernel, $\mathsf{Lin} \times \mathsf{SE}$.

6.4 | Training the model

When defining the covariance function for a GP it may be desirable to specify a number of free hyperparameters, θ, which allow the properties

Kernel	Symbol	Properties
Exponential-quadratic	SE	C^{∞}-smooth local variation.
Matérn-3/2	M32	C^{3}-smooth local-variation
Matérn-5/2	M52	C^{5}-smooth local-variation.
Periodic	Per	Smooth global periodic variation.
Linear	Lin	Global continuous linear variation.
Rational Quadratic	RQ	Variation on multiple scales.
Constant	C	Scaling factor.

Table 6.3.1: Frequently used and encountered kernels used as covariance functions for GPR problems. The second column contains the abbreviation by which these kernels are referred in this work, and the third column lists properties of each function which affect its utility in a variety of problems.

of the GP to be altered. Since the functional form of the covariance function defines the GP model, this allows the techniques of Bayesian model selection to be employed, in order to select the specific GP model which optimally describes the data. The log-probability that a given set of strain values were drawn from a Gaussian process with zero mean and a covariance matrix $\mathsf{K} = K_{ij} = k(x, x'; \theta)$ is

$$\log p(\boldsymbol{f}|x) = -\frac{1}{2}\mathsf{K}^{-1}\boldsymbol{f} - \frac{1}{2}\log|\mathsf{K}| - \frac{n}{2}\log 2\pi. \tag{6.18}$$

This quantity is normally denoted the *log-evidence* or the *log-hyperlikelihood*. The model which best describes the training data may then be found by maximising the log-hyperlikelihood with respect to the hyperparameters, θ of the covariance function, $k(x, x'; \theta)$.

This optimisation may be conducted using either a hill-climbing based optimisation algorithm, or in a hierarchical Bayesian framework, with prior probability distributions assigned to each hyperparameter, and the optimal hyperparameters then found using an Markov chain Monte Carlo (MCMC) algorithm.

6.5 | Dealing with computational complexity and large data sets

One severe disadvantage of Gaussian Processes as a data analysis tool are their high computational complexity. Producing a prediction from a GP requires inverting the covariance matrix; matrix inversion is an $\mathcal{O}(N^3)$ process in time, and scales with $\mathcal{O}(N^2)$ in memory use. This effectively limits the number of training points which can be input to a GP to fewer than 10^4.

A number of approaches have been developed in the literature to address this short-coming by utilising computationally tractable approximations to either the matrix inversion or the Gaussian process. These approaches can be grouped into three broad categories; sparse Gaussian processes, which use a modified covariance function to force the covariance matrix to have a near-diagonal structure; hierarchical approaches, which do not modify the covariance function, but approximate the off-diagonal terms' influence on the inversion; and local expert approaches, in which the parameter space is divided into many sub-spaces, and each sub-space is modelled using an independent GP.

6.5.1 Sparse Gaussian processes

Sparse GPR approaches work by modifying the form of the joint prior distribution from equation 6.5 to include an additional m latent variables,

$$\boldsymbol{u} = [u_1, \ldots, u_m]^\intercal,$$

which are termed "inducing variables". These correspond to values of the Gaussian process at inputs $X_{\boldsymbol{u}}$, which are the inducing inputs. These inducing variables can be chosen in various different ways, but their effect on the GP is the same.

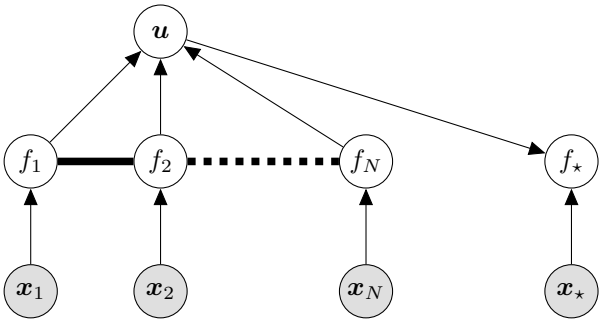


Figure 6.5.1: A graphical model of a sparse Gaussian process, represented as a chain graph. The inputs (on the bottom row) are all observed quantities. For the sake of clarity the outputs have been omitted from this diagram. The latent variables, f from the Gaussian field (the heavy black line connecting these nodes indicates that they are fully connected) connect the two, and so any given observation is independent of all other nodes given it connected latent f variable. In contrast to the fully-connected situation depicted in 6.2.2, the values of the Gaussian process for the training data are taken to be conditionally independent from the values for test inputs.

The original GP can be recovered by marginalising over $\boldsymbol{u}$:

$$p(\boldsymbol{f}_*, \boldsymbol{f}) = \int p(\boldsymbol{f}_*, \boldsymbol{f}, \boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} = \int p(\boldsymbol{f}_*, \boldsymbol{f}|\boldsymbol{u}) p(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \tag{6.19}$$

with $p(\boldsymbol{u}) = \mathcal{N}(\mathbf{0}, \mathsf{K}_{\boldsymbol{u},\boldsymbol{u}})$.

Sparse GP approaches make the assumption that $\boldsymbol{f}$ and $\boldsymbol{f}_*$ are conditionally independent given $\boldsymbol{u}$. This is depicted as a graphical model in figure 6.5.1.

This allows the construction of two conditional posterior probability distributions, for the training data and the test inputs [260]:

training:

$$p(\boldsymbol{f}|\boldsymbol{u}) = \mathcal{N}(\mathsf{K}_{\boldsymbol{f},\boldsymbol{u}}\mathsf{K}_{\boldsymbol{u},\boldsymbol{u}}^{-1}\boldsymbol{u}, \mathsf{K}_{\boldsymbol{f},\boldsymbol{f}} - Q_{\boldsymbol{f},\boldsymbol{f}}) \tag{6.20a}$$

test (predictive):

$$p(\boldsymbol{f}_*|\boldsymbol{u}) = \mathcal{N}(\mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{u}}\mathsf{K}_{\boldsymbol{u},\boldsymbol{u}}^{-1}\boldsymbol{u}, \mathsf{K}_{\boldsymbol{f}_*,\boldsymbol{f}_*} - Q_{\boldsymbol{f}_*,\boldsymbol{f}_*}) \tag{6.20b}$$

letting $Q_{\boldsymbol{a},\boldsymbol{b}} = \mathsf{K}_{\boldsymbol{a},\boldsymbol{u}}\mathsf{K}_{\boldsymbol{u},\boldsymbol{u}}^{-1}\mathsf{K}_{\boldsymbol{u},\boldsymbol{b}}$.

There are a number of approaches to choosing the inducing points, and further simplifying assumptions which can be applied to the sparse GP approach which are discussed in depth in [260]. Thanks to the smaller matrix which must be inverted for the predictive case, formed only from the inducing points, this sparse approach is capable of handling much larger quantities of data than the direct, exact approach.

6.5.2 Hierarchical matrix solvers

An alternative approach to introducing an inducing set is to take advantage of the structure of the covariance matrix, K, which is produced by a number of covariance functions. Covariance functions will typically assign a small covariance to points which are distantly spaced in the data space; as a result, if the covariance matrix is suitably sorted, it is possible to conside the whole covariance matrix as a block matrix. Hierarchical solving methods such as [261, 262] produce an arrangement of low-rank matrices as off-diagonal components in the block matrix. The on-diagonal sub-matrices are still treated as full rank matrices, and are solved using conventional methods, while the inverses of the off-diagonal components are found using a Chebyshev polynomial interpolation and LU-decomposition. This allows for inversion of the matrix in $\mathcal{O}(n \log^2 n)$ rather than $\mathcal{O}(n^3)$ time. This technique has been successfully applied to GPs in the `George` library [263].

6.5.3 Gaussian process local experts

Local expert approaches attempt to improve the computational performance of GPs by diving the parameter space of the model into multiple sub-spaces. In a conventional GP the training data is used in its entirity to train a single GP. If these data were instead divided into M subsets, of

size K, we can train M separate GPs, which will each provide an independent prediction for any given point in the parameter space. The network structure which is established by this subdivision of the parameter space is known as a *gating network*.

Early approaches to using local experts in GPs used kd-trees [264] to sub-divide the parameter space, and then modelled each subspace with its own GP. The GPs were trained together, with each having the same kernel hyper-parameters. Final predictions were then produced as a weighted sum of the individual GPs' predictions. While this approach was somewhat effective, it enforced a stationary structure on the covariance matrix, and the paper does not treat the combination of the prediction uncertainties.

Approaches which follow the work of [265] on mixtures of local experts have had some more promise, allowing each GP to have its own set of hyper-parameters, allowing greater freedom in modelling heteroskedastic and non-stationary data.

Deciding on the number of sub-models is a non-trivial problem; one approach is to model the parameter space using an infinite mixture model (IMM) [266], in which the gating network is effectively a Dirichlet process over the training data. The predictions from each sub-model are then summed to find the global prediction. While this approach offers greater flexibility for modelling more complex underlying functions, it does little to improve the speed of GP predictions. Additional IMM approaches are proposed by [267], and a comparable, variational approach is taken by [268].

All of these approaches have the difficulty of requiring the gating network to assign a weight (often called a *responsibility* to each sub-model's prediction when calculating the global prediction, adding an additional layer of inference, which normally requires an MCMC sampler to perform. *Product-of-experts* models avoid this complication by multiplying the sub-

model predictions, but these models have either turned out to be excessively confident [269], or excessively conservative [270].

These problems have lead to the development of the Bayesian Committee Machine (BCM) [271], which assigns a weight to each sub-model's prediction which is equal to the inverse of the prediction's covariance, in order that sub-models which better observe the predicted region are given a greater weight in the global prediction. This approach can suffer as a result of models which contains week experts, and so the *robust Bayesian Committee Machine* [272] has been proposed to provide a more robust framework for Gaussian process regression with many experts. This approach also allows for the computation of the model's prediction to be highly-parallelised, with the potential for each sub-model being evaluated on separate compute nodes, and combined together by another process running on another node.

6.5.4 Stochastic Variational Inference

The stochastic variational inference (SVI) algorithm is designed to allow inference to be carried out in situations where very large quantities of data are available.

Variational inference, whereby a posterior distribution over some set of latent variables $\mathcal{Z}$, given data $\mathcal{D}$ is approximated with a *variational distribution*:

$$P(\mathcal{Z}|\mathcal{D}) \approx Q(\mathcal{Z}) \tag{6.21}$$

where the distribution $Q(\mathcal{Z})$ is restricted to be simpler than the form of the exact posterior. The similarity between Q and P can be measured with the Kullback-Liebler divergence (see definition 4.2.4); as such, finding a suitable approximation of the posterior distribution becomes a standard optimisation problem, in which the KL divergence must be minimised.

Stochastic optimisation is designed to find the maximum of an objective function by following noisy estimates of the function's gradient; these gradients must be unbiased. Variational inference has the attractive property that the objective function can be decomposed into additive terms, with one term for each datum in $\mathcal{D}$. Noisy estimates of the gradient can be obtained by taking a subsample of $\mathcal{D}$ and using it to compute a scaled gradient on that subsample. If sampled independently the gradient of the noisy gradient will be equal to the true gradient [273].

This combination of stochastic optimisation and variational inference is suitable for models which have a set of global variables which factorise the observable and latent variables of the model, however, the graphical model of a GP, as depicted in 6.2.2 makes it clear that these models do not possess such a structure. However, *sparse* GP models do possess a structure with global variables, thanks to the existence of the set of inducing points. The structure of these models, depicted in figure 6.5.1 is close to the requirement for SVI, as the global variables factorise the observable variables.

For a GP model to use SVI a variational distribution is introduced over the inducing variables: $q(\boldsymbol{u})$. This distribution is Gaussian, and can be parameterised as $q(\boldsymbol{u}) = \mathcal{N}(\boldsymbol{u}|\boldsymbol{m}, \boldsymbol{S})$. A lower bound can be set on the distribution (see equation 4 of [274]) by Jensen's inequality. This lower bound can be expressed as a sum of terms which correspond to single pairs $(\boldsymbol{x}, y)$ from the training set, which allows stochastic optimisation to be carried-out.

The use of a posterior approximated by variational inference in this way allows for much larger datasets to be used in the conditioning of the GP than other methods, since only a subset (or "minibatch" of the training data must be used in any given training iteration).

6.6 | Assessing Gaussian process regression models

Having produced a statistical regression model it is crucial that its efficacy is assessed. There are broadly two scenarios under which such testing can occur. In situations where a large amount of data is available to condition the model it is often appropriate to partition the data into a "training set" and a "test set"; the latter is held-aside, and not used to condition the model, and can then be used after the model is trained to compare against the model predictions.

Alternatively scenarios may arise where there is insufficient data to form such a test set without adversely affecting the model's predictive power. Examples of such a scenario include timeseries modelling, where the predictions of the model may represent future (an therefore inaccessible) observations, or computational experiments, where the acquisition of training data is sufficiently costly that producing a test set is not viable.

In the case where test data is available two straight-forward metrics are available: the root-mean-squared error, and the correlation.

Let $\boldsymbol{x}_*$ and $\boldsymbol{y}_*$ be respectively the test inputs and test outputs from the test set, then let $\hat{\boldsymbol{y}}$ be the set of model predictions drawn from the Gaussian Process with inputs $\boldsymbol{x}_*$.

The root mean squared error (RMSE) gives an estimate of the total deviation between the mean prediction of the model and the true value from the test data:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{n_i} (y_*^{(i)} - \hat{y}^{(i)})^2}{n_t}}, \tag{6.22}$$

for n_t the size of the test set. While the RMSE can represent a good metric for conventional regression methods, it does not consider the estimate of

the variance which is provided by Gaussian process models; as such it is an insufficient measure on its own of these models.

It is possible to use the GP variance to form a metric of the efficacy by considering the correlation between the test data and the prediction

$$\rho^2 = \left(\frac{\operatorname{cov}(y^*, \hat{y})}{\sqrt{\operatorname{var}(y) \operatorname{var}(\hat{y})}} \right)^2 \tag{6.23}$$

These two metrics, together, allow the model to be assessed either during the training of the model (or indeed, they can be used as training metrics if using a cross validation approach while determining the model hyperparameters) given a judicious partitioning of the available data.

Forrester [275] suggests that a $\rho^2 \geq 0.8$ provides a surrogate model with good global predictive abilities, which corresponds to an RMSE of around 0.1.

In situations where test data is not available such straightforward tests are often impractical. In the case of timeseries forecasting it may be possible to assess the forecast by forming a test set from the most recent observations, and comparing these to the output of the model, however, if only a small number of past observations are available the predictive capability of the model may be sufficiently poor to render this test almost meaningless.

In situations where more data is available it may be possible to assess a GPR model using *leave-one-out* cross validation, in which a single point is omitted from the training set, and used as test data. The testing can then be repeated multiple times, leaving different points from the sample in order to form a comprehensive test statistic.

6.7 | Estimating contours: an example GPR problem

While figure 6.2.1 showed the process of constructing a GP regressor for data generated from a single-dimensional function, in this section I demonstrate how a higher-dimensional problem can be treated with GP regression. For the sake of clarity I have chosen a two-dimensional function; anything with more dimensions is likely to be hard to represent on paper, and the same concepts can be extended to higher-dimensional models.

In figure 6.7.1 a number of spot-heights are plotted for hills in the *Arrochar Alps*, a region of the Scottish Highlands around 50-kilometres north of the City of Glasgow. Each point corresponds to the summit of a hill (derived from the *Database of British and Irish Hills* [276]). In order to interpolate a "landscape" based on these measurements I trained a GP with a rational quadratic kernel on the latitude and the longitude. The α parameter of the kernel was set to be the same in both dimensions, and a Γ-function prior was placed on it with shape parameters $(\alpha_\Gamma = 5, \beta_\Gamma = 0.5)$. A normal distribution prior was placed on the lengthscale of each dimension, each with $(\mu = 0.012, \sigma = 1)$. It is worth noting that applying a constraint on a GP is difficult, and as such, despite providing the peak heights in the landscape, the GP is free to interpolate larger height values throughout the landscape. Finally, the covariance function was multiplied by a constant kernel scaling factor (C), the amplitude of which was drawn from a normal distribution prior with parameters $(\mu = 1, \sigma = 1)$.

The GP was implemented using the `PyMC3` python library [250].

In order to determine the appropriate hyperparameter values the log-evidence was maximised using a Newtonian optimiser, in order to deter-

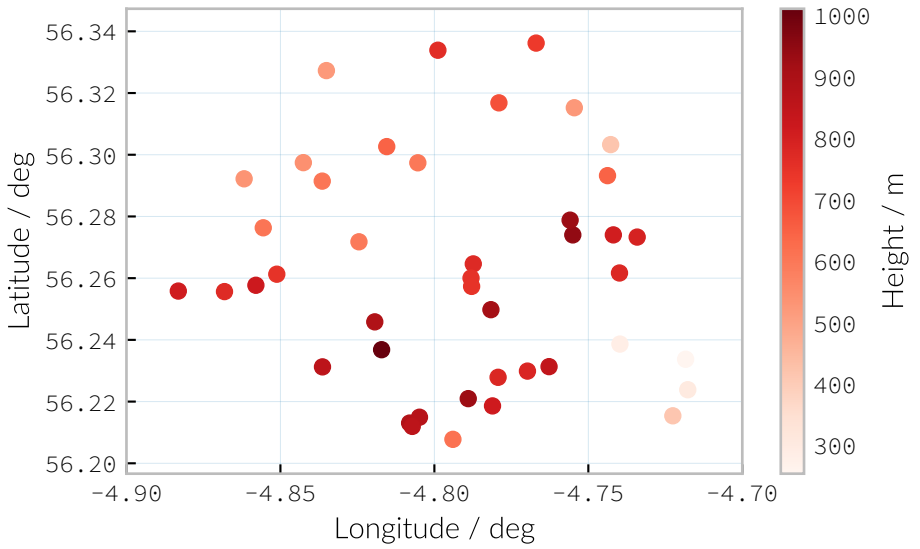


Figure 6.7.1: The location of summits within the *Arrochar Alps*, an uplands region of Western Scotland. These will be used as the training data for a GP regression model designed to emulate the landscape.

mine the maximum a posteriori estimate (MAP) estimate of the hyperparameters. The resulting MAP estimate of the mean landscape is shown in figure 6.7.2. A number of *irregularities* can be spotted with a map produced using this technique, rather than a more standard method. The first is the absence of a flat region of land occupied by a large reservoir between *Ben Vane* $(56.249786°, -4.781639°)$ and *Ben Vorlich* $(56.274021°, -4.755046°)$; as the map is informed only by summits this surrogate model for the landscape is bound to struggle to find low points like this in the landscape. The second is the very smooth nature of the landscape, for example the near-conical shape of *Beinn Ìme* $(56.236812°, -4.817142°)$; this is a result of the choice of a smooth kernel (the RQ kernel). The behaviour of the GP far from any of the training data is mostly obscured in this figure thanks to the clipping of the boundary box; the GP will eventually revert to the mean of the GP prior (which was chosen to be zero in this example); this

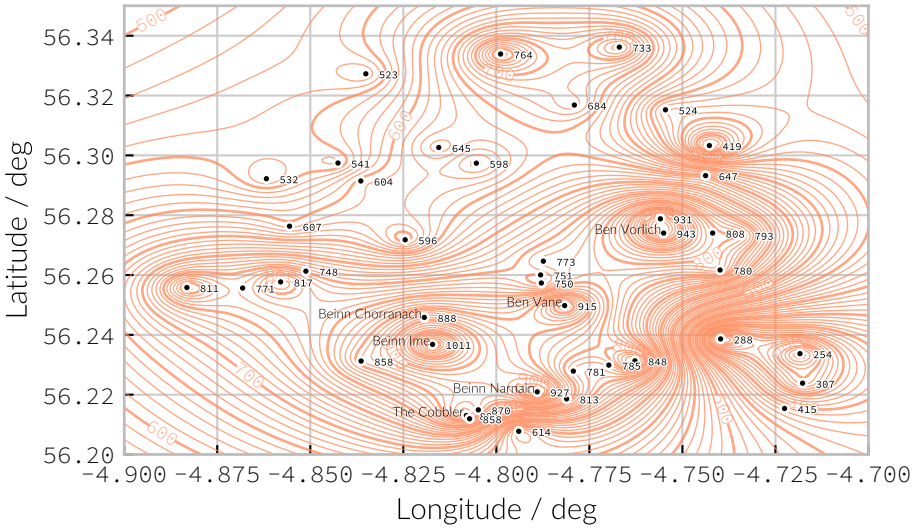


Figure 6.7.2: The mean GP output for a GP trained with summit heights in the Arrochar Alps, an upland area north of Glasgow, Scotland. Here the smoothness conditions placed on the GP by the form of the covariance function become clear with a number of the peaks being lost as a result. In this example a rational-quadratic covariance function was used.

behaviour can be seen to some extent in the upper-left corner of the plot.

In figure 6.7.3 I show the same landscape created using GPs with a variety of covariance functions which show how drastically this choice affects the model.

Four different covariance functions are shown; constructed from the rational quadratic (RQ), Matérn-5/2 (M52), exponential quadratic (SE), and the exponential kernels respectively. The variance of the predictions from each GP are shown in figure 6.7.4.

Each of these predictions show behaviour created by the choice of covariance function. The rational quadratic covariance function infers a smooth, rolling landscape between the peaks, but still produces pronounced peaks. The prediction with this covariance kernel is confident throughout the area of the plot, as seen from the low variance in the

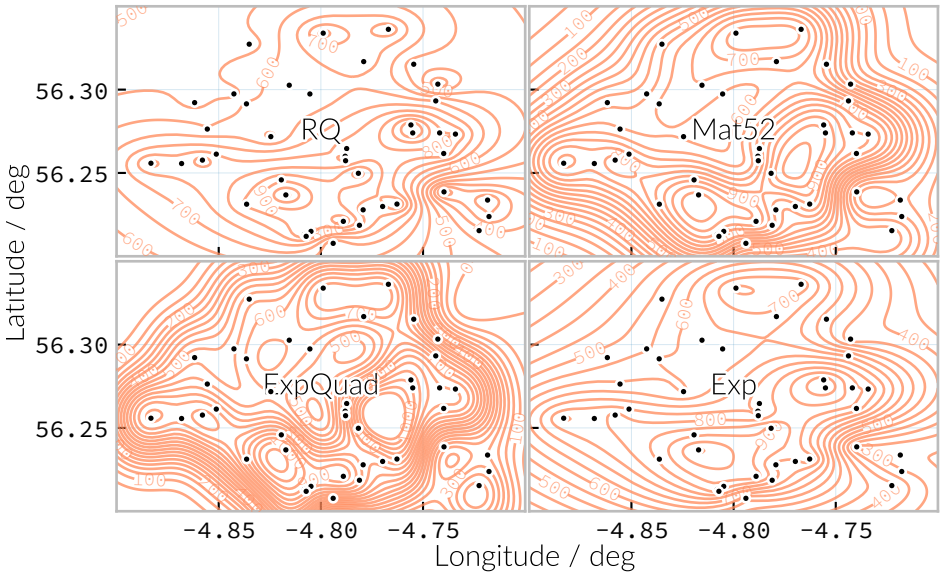


Figure 6.7.3: The GP derived mean landscape, with a variety of different covariance functions used to produce the interpolated topology. The upper-left panel is generated from a GP with a rational quadratic kernel (this is a repeat of figure 6.7.2); then the upper right is generated using a Matérn-5/2 kernel, lower left an exponential quadratic kernel, and lower right an exponential kernel. Each panel also contains the training points marked as black dots.

upper-left panel of figure 6.7.4.

The behaviour of the GPs which use Matérn-5/2 and exponential quadratic covariance function are broadly comparable, favouring much steeper slopes than the rational quadratic GP, and providing low-confidence predictions in regions outside the training data. This effect is moderately more pronounced for the GP using the exponential quadratic than the Matérn-5/2.

Similarly to the rational quadratic kernel, the GP using the exponential kernel produces a landscape with smoothly-varying large-scale structure, but allows for steeper gradients close to training points, and produces lower-confidence estimates than the GP using the rational quadratic covariance function outside of the parameter space spanned by the training data.

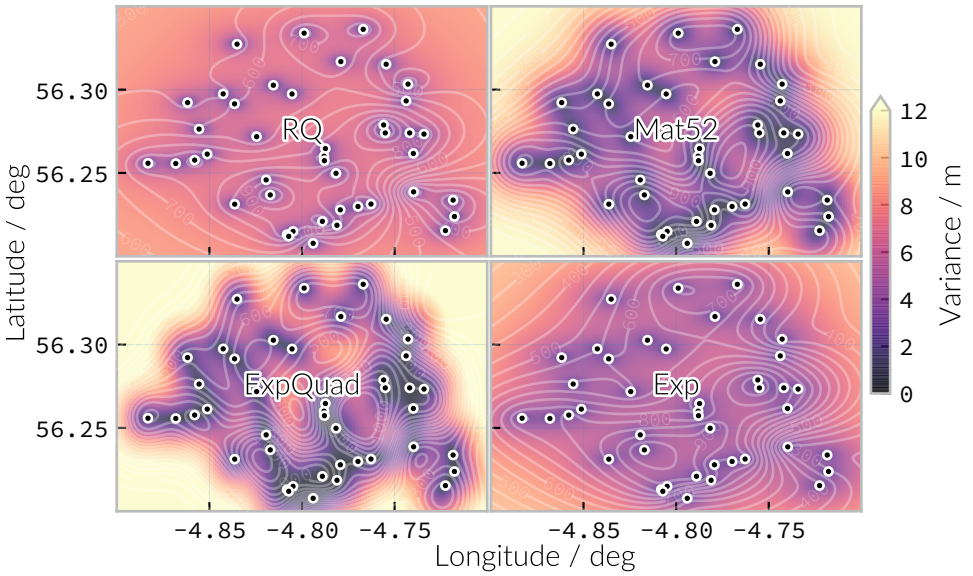


Figure 6.7.4: The variance of the landscapes from figure 6.7.3, with the uncertainty underlaid as a colourmap, which runs from dark in regions of low variance, generally close to the peaks, where the training data was provided to the GP, to light in regions of high variance (and hence high uncertainty).

While this is clearly not a practical method for use in cartography, the behaviour of the four GPs shown in figures 6.7.3 and 6.7.4 is helpful to understand the behaviour of GP in higher-dimensional spaces.

7 Heron: A Gaussian process regression approach to modelling gravitational waveforms

In this chapter I introduce the use of Gaussian process regression as a surrogate modelling method for gravitational wave (GW) waveforms from compact binary coalescence events. Some of the work presented in this chapter has previously been made available on the ArXiv [2], and at the time of writing are under review for publication. The culmination of this effort is the development of the *Heron* surrogate model, which is capable of producing waveforms with precession for arbitrarily spinning binary systems. Additional results are presented, specifically related to the comparison between the Heron model and preexisting approximate models.

In section 7.1 I discuss the training data used for the model, which is derived from the Georgia Tech waveform catalogue, and how the waveform data from the catalogue are prepared. In section 7.2 I extend the discussion of Gaussian process regression from chapter 6 to the case of gravitational waveform data, and produce two simplified example waveform models using this technique. In section 7.3 I describe the development of the

Heron model itself, using numerical relativity data, and discuss some of the practical computational considerations for a model using this data. In section 7.4 I discuss a number of tests which can be performed on the Heron model in order to demonstrate its efficacy across its parameter space, and I present results from these tests. In section 7.5 I plot a number of example waveforms produced by the Heron model. Finally, section 7.6 provides an overview of the material in the chapter, and some discussion of future work which would be of value to the continued development of this model.

This work was made possible thanks to data provided by Georgia Institute for Technology (Georgia Tech) and the Simulating Extreme Spacetimes (SXS) collaboration.

The production of high-accuracy waveforms is possible thanks to advances in the field of numerical relativity (NR), in which the full set of Einstein equations can be solved numerically (see section 2.3.3 for a more detailed discussion of the development of NR). This can be done reliably for the low-mass compact binary systems of interest to the current generation of ground-based gravitational wave observatories, however these simulations are computationally expensive, and can require thousands of CPU hours to run in situations where the mass ratios and spins of the black holes are small. A simulation of a full 350-cycle gravitational waveform spanning the entire advanced Laser interferometer gravitational-wave observatory (LIGO) band has been produced [277], however this required several months of high-performance computing to complete [278], despite employing numerous techniques to reduce wall-clock computation time. As a result fewer than 1000 waveforms are available, many of these much shorter than 350 cycles long. Binary black hole (BBH) coalescences are described by a number of physical parameters: the ratio of the two component black holes' masses, q; the vector of each component's spin, $\boldsymbol{s}_1 = (s_{1x}, s_{1y}, s_{1z})$ and $\boldsymbol{s}_2 = (s_{2x}, s_{2y}, s_{2z})$, defined such that the z-component of each vector is parallel to the total angular momentum vector, $\boldsymbol{L}$, for the system (i.e. $\hat{s}_{1z} = \hat{s}_{2z} = \hat{L}$, see figure 7.0.1 for the arrangement of these vectors); and the time, t, relative to a fixed reference time, for example the time of coalescence of the binary.

As noted in section 2.3.5, a number of *analytical approximant* waveforms have been produced in recent years in order to provide approximate waveforms much more quickly than NR is capable of. These models, while capable of producing waveforms over a wide range of the BBH parameter space, are calibrated against only a small number of waveforms, resulting in waveforms produced in much of the parameter space having an unknown accuracy.

7.1 | Numerical Relativity Waveform data

NR simulations for compact binary coalescence (CBC) events simulate the behaviour of the Riemann tensor over a finite period of time prior to, during, and immediately following the merger of the two objects. This simulation allows the extraction of a gravitational waveform at a large distance from the merger event, where the flat metric approximation of equation 1.2 is valid. Observations of the strain are made, and time-stamped. These data then constitute the waveform data, which are normally stored in a spherical harmonic decomposition in an `HDF5` format [279]. Thanks to the non-uniform time sampling which NR codes employ, the data from the simulation will normally need to be interpolated (in time) and resampled to be of use for data analysis, or as training data for a surrogate model. In the models presented in this chapter this interpolation was performed by the `pycbc` library [242, 243, 280].

The waveform data used in this chapter is derived from the Georgia Tech Catalogue [157], with the waveforms evaluated in the time domain,

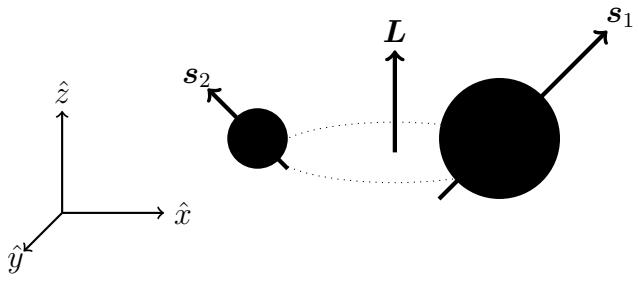


Figure 7.0.1: A cartoon depicting the geometrical arrangement of the two black holes in a coalescing binary system, and the various vector quantities involved. Here $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are the spin vectors of the individual component black holes during the inspiral, and $\boldsymbol{L}$ is the total angular momentum vector for the system. The spin vectors are defined such that $\hat{s}_z$ for each is equal to $\hat{L}$, and the other two components are orthogonal.

and then aligned such that the time location in the timeseries corresponding to the maximum value of root square sum strain (hrss) was labelled 0. A timeseries is generated for each GW polarisation, giving two timeseries, for $h_+(t)$ and $h_\times(t)$ for each NR waveform.

Each waveform is parameterised by seven quantities (the mass ratio and the spin vectors of each component black hole) in a vector which is denoted $\boldsymbol{x}_i$. Each strain value, h_i, within the waveform is further parameterised by a time relative to the maximum strain value in the waveform, and thus each training point is parameterised by an 8-dimensional parameter vector, denoted $\boldsymbol{x}_i{}'$. This provides a training set which has 8 input dimensions, and a single output dimension, with the form

$$\mathcal{D} = \{(\boldsymbol{x}_i', h_i) | i = 1, 2, \ldots, N\} \tag{7.1}$$

for N the total number of strain samples used from all of the training waveforms, and $h_i = h(t_i, \boldsymbol{x}_i) = h(\boldsymbol{x}_i')$ is the strain value of a waveform at time t_i relative to the maximum strain value of the waveform, and at a location $\boldsymbol{x}$ in parameter space. The distribution of training waveforms throughout the parameter space is shown in figure 2.3.3, on page 64.

7.2 | Gaussian process regression for waveforms

Current waveform models lack a natural means to account reliably for uncertainty in the waveforms which they generate. A side-effect of this is important in parameter estimation of these signals, as systematic uncertainties in the waveform model will result in the introduction of a bias to the parameter estimation process.

To illustrate how such a model can be constructed, in this section I will detail the construction of two Gaussian process (GP) models which use

waveforms generated from the `IMRPhenomPv2` waveform model as their training data. In the first model I will use only non-spinning waveforms, varying only in their mass-ratio. In the second I will introduce characteristic spin to the model, to create a three-dimensional GP regression model.

Later in this chapter I will detail the results presented in [2] of a GP model for precessing waveforms which has been trained using data from NR simulations.

7.2.1 A simple non-spinning model

For the first demonstration model I will use 15 waveforms produced using the `IMRPhenomPv2` waveform model with mass-ratios spaced uniformly in log-space between $q = 0.1$ and $q = 1.0$. Each waveform is composed only from the $(2, 2)$ spherical harmonic mode, and the strain is labelled as $h_{2,2}$. Each waveform was sampled at $4096\,\mathrm{Hz}$ for a total system mass of $60\,\mathrm{M}_\odot$. The strain data is generated assuming that the source is placed at a distance of $1\,\mathrm{M}_\odot$ (or approximately $1.477\,\mathrm{km}$) from the observer, allowing the strain to be easily rescaled to any distance, r, by dividing the amplitude by the distance to the source (in units of $\mathrm{M}_\odot$). The strain data used for training is then denoted $rh_{2,2}$. The timestamps are then divided by the total mass of the system to produce a waveform which can be rescaled for an arbitrary total mass by multiplying the timestamps by the desired total mass (in units of $\mathrm{M}_\odot$). The training data is plotted in figure 7.2.1.

The covariance structure for this model consists of two squared exponential (SE) kernels (one for each dimension, time and the natural logarithm of the mass ratio) and a constant kernel as a scale-factor:

$$K = \mathsf{C} \times \mathsf{SE}^{(t,\log q)}. \tag{7.2}$$

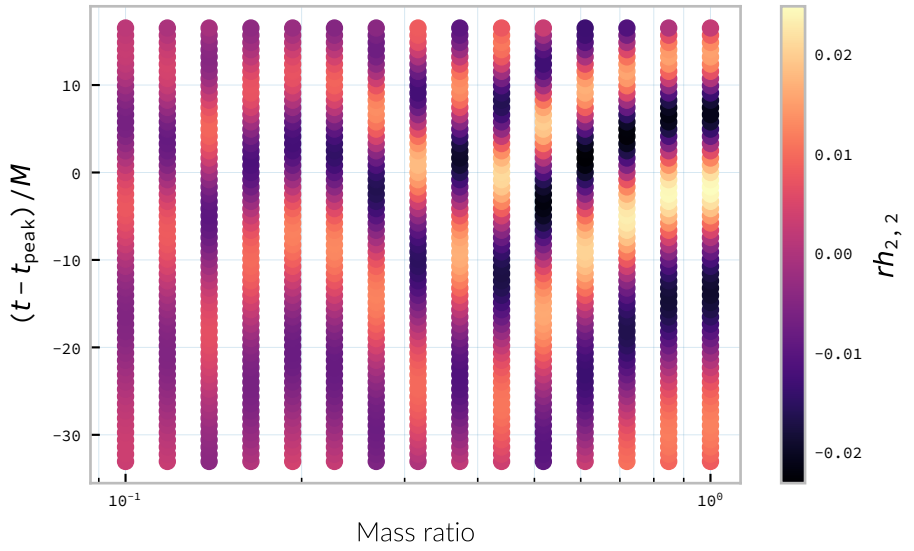


Figure 7.2.1: The training data used for the non-spinning demonstration GP surrogate model. Here the amplitude of the data, $rh_{2,2}$ represents the $(2, 2)$-mode strain, multiplied by the distance from the source to the observer. The physical strain can be recovered by dividing this quantity by the desired distance from the observer to the source.

SE covariance functions were chosen as it allows smooth data generating functions to be modelled well, and there was no reason to assume that the function contained discontinuities, or non-smooth regions. The constant kernel is included to allow the GP extra flexibility to determine the variance of the estimate. A small amount of noise is added to the covariance function (with a mean of 1×10^{-6}) to improve the numerical stability of the matrix inversion process. The hyperparameter values for each part of the kernel were determined by maximising the log-likelihood of the GP to complete the training of the model.

A sample of a waveform generated from this model is plotted in figure 7.2.2, with the mean and variance of the GP plotted as a dashed grey line and shaded grey region respectively. Individual draws from the predictive posterior distribution are plotted as solid grey lines, depicting how

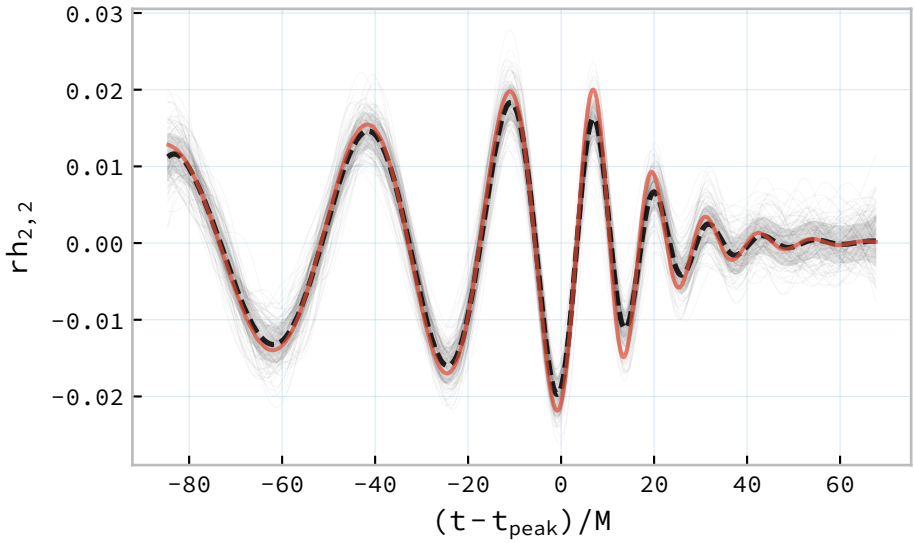


Figure 7.2.2: One hundred draws from a GP trained on waveform data produced from the `IMRPhenomPv2` analytical waveform model. These draws are produced for a non-spinning, non-equal-mass configuration ($\boldsymbol{s_1} = (0, 0, 0)$, $\boldsymbol{s_2} = (0, 0, 0)$, $\boldsymbol{q} = 0.8$), and each is shown shown as a light grey line. The output of `IMRPhenomPv2` is overlaid in red. The mean draw from the GP is shown as a grey dashed line, while the associated variance is plotted as a grey-filled region surrounding the mean.

individual fitting functions drawn from the posterior behave, clustering closely about the mean, with the behaviour in the ringdown portion of the waveform showing the greatest variance, which is reflected in the greater width of the variance region in this part of the waveform.

As this is a two-dimensional model, it is possible to inspect the entire surface of the mean function produced by the surrogate model; this, along with the variance of the model prediction is plotted in figure 7.2.3.

In order to compare the output of the GP surrogate model to `IMRPhenomPv2` I calculate the match of the waveform generated by each model. The match is defined between two waveforms as

$$\mathcal{M}(h_{\text{model}}, h_{\text{ana}}) = \max_{t_0, \phi_0} \frac{\langle h_{\text{model}}, h_{\text{ana}} \rangle}{\sqrt{\langle h_{\text{model}}, h_{\text{model}} \rangle \langle h_{\text{ana}}, h_{\text{ana}} \rangle}}. \tag{7.3}$$

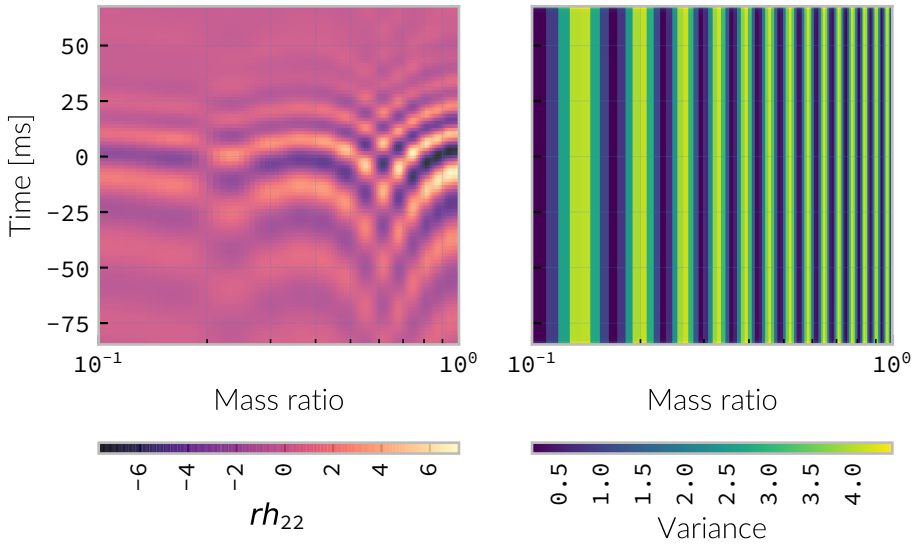


Figure 7.2.3: The mean (left panel) and variance (right panel) of the GP surrogate model trained on `IMRPhenomPv2` over the (t, q)-plane.

where h_{model} and h_{ana} are respectively the timeseries predicted by the model and the analytical approximant, t_0 and ϕ_0 are the merger time and merger phase, and $\langle \cdot, \cdot \rangle$ is the noise-weighted inner product between two waveforms. This is defined as

$$\langle a, b \rangle = \Re \int_{-\infty}^{\infty} \frac{\tilde{a}^*(f)\tilde{b}(f)}{S_n(f)} \mathrm{d}f \tag{7.4}$$

for $\tilde{a}$ and $\tilde{b}$ respectively the Fourier transforms of the timeseries a and b, S_n the amplitude spectral density of the noise, and f the frequency. It will also be convenient at this point to define the *mismatch* as $1 - \mathcal{M}$.

The match is a useful quantity to consider, as it is widely used in GW signal analysis, thanks to its use in matched filtering, which is the optimal filter for maximising signal-to-noise ratio (SNR) when additive noise is present in data [15].

I calculate the (noise-free, with $S_n(f) = 1$) match between the mean waveform from the GP surrogate and the waveform from `IMRPhenomPv2` for 100 mass ratios between $q = 0.1$ and $q = 1.0$ in figure 7.2.4. The

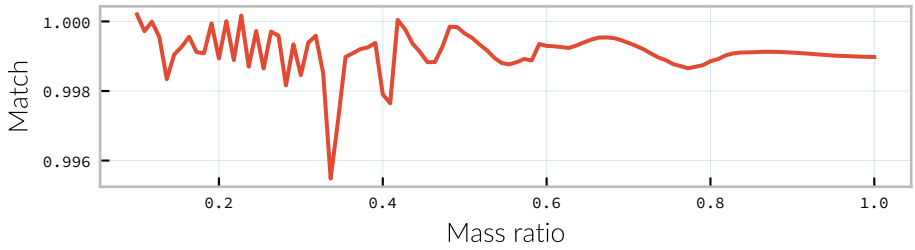


Figure 7.2.4: The match between the waveforms produced by the `IMRPhenomPv2` and GP models across one hundred different mass ratios. The mean waveform from the GP model was used for this plot.

match of the model is consistently above 0.995 at each mass-ratio tested (although there is a notable drop in match around 0.35 corresponding to one of the locations furthest from training data). While the use of match demonstrates that the GP is capable of reproducing the structure of a waveform accurately it will fail to characterise problems with a model which incorrectly predicts the amplitude of the waveform. The match (and mismatch) are the standard metric which is used for evaluating waveform models, however there may be need for additional checks (beyond simply verification by-eye as in figure 7.2.2).

7.2.2 A model with effective spin

In order to extend the model from two dimensions (time and mass ratio), in this section I will introduce a model which additionally models systems with spin. For simplicity this spin is limited to parallel-spinning systems in which each component black hole has the same spin. Additionally, for this demonstration, all of the waveforms used in training were from systems where the spin of the component black holes was parallel to the total angular momentum of the system.

As with the non-spinning model in the previous section, the covariance

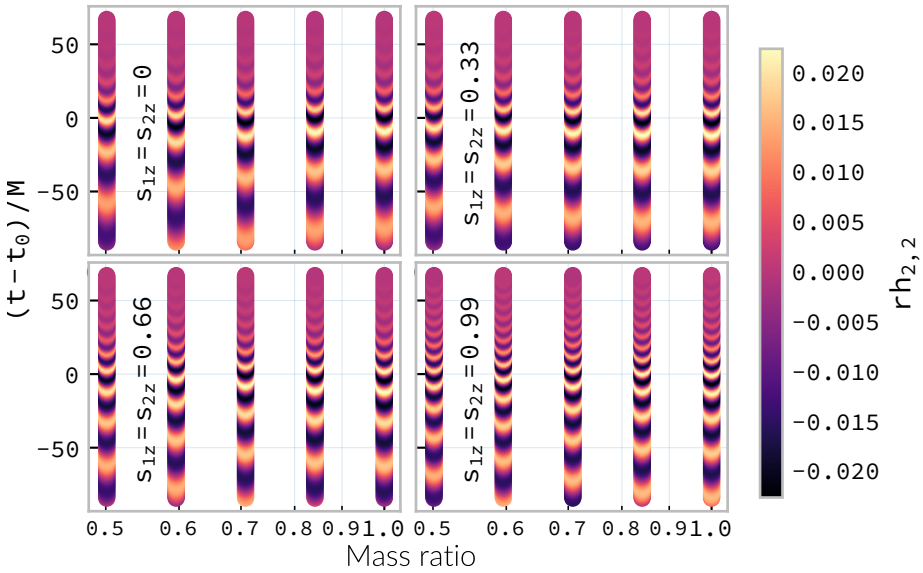


Figure 7.2.5: The training data for the spinning GP surrogate model, derived from the `IMRPhenomPv2` approximant model. Four different values of spin were used to produce this data, $s_{1z} = s_{2z} = \{0, 0.33, 0.66, 0.99\}$.

structure for this model consists of squared exponential kernels, with two additional dimensions added for the z-component of each black hole's spin:

$$K = \mathsf{C} \times \mathsf{SE}^{(t, \log q, s_{1z}, s_{2z})}. \tag{7.5}$$

As with the non-spinning model in the previous section the SE kernel is used to model each data-space dimension as these are expected to vary smoothly.

Again the model is trained with a Newtonian optimiser, and tested against the direct output of the `IMRPhenomPv2` model, by calculating the noise-free match over the parameter space. The results of this comparison are displayed as the surface plot in figure 7.2.6. In the majority of the parameter space the mismatch is very small (less than 0.05), however the GP model struggles to replicate the behaviour of `IMRPhenomPv2` at high-spins (where the waveform evolves rapidly due to the extreme physics involved), and at low mass-ratios (corresponding to asymmetric systems)

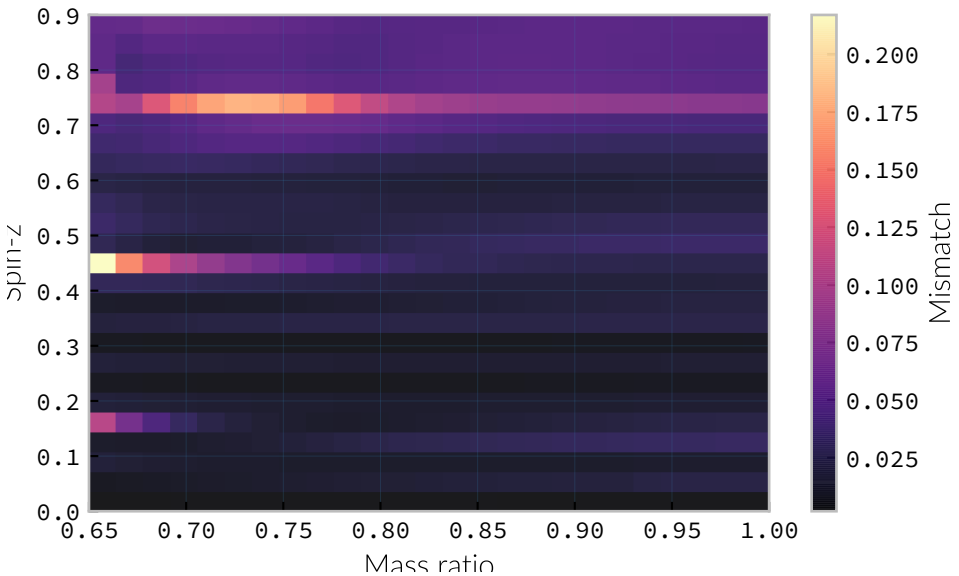


Figure 7.2.6: Matches across the mass-ratio and $s_{1z} = s_{2z}$ plane between the GPR model and waveforms generated directly from `IMRPhenomPv2`.

where again the behaviour of the waveform is more complicated thanks to the underlying physics of the orbit. Three bands of high mismatch can be seen at locations approximately half-way between the locations of waveform samples in the spin dimension, implying that more than four samples are required for a better model.

Using an analytical approximant model such as `IMRhenomPv2`, which was used here, or `SEOBNRv3`, grants the ability to generate training data on demand and at whim; this is a situation which is not possible when using NR waveforms. In the following sections I will discuss the development of a model trained entirely off these waveforms, and the challenges which this approach posed.

7.3 | Heron: A model using numerical relativity waveforms

The Heron model was created by training a GP regression model on data from the Georgia Tech BBH waveform catalogue [157]. It is designed as a proof-of-concept surrogate model which operates over the eight dimensions of the BBH parameter space, demonstrating both that GP regression is a useful technique for dealing with this high-dimensional problem, but also that training can be carried-out directly from NR data.

The model is constructed using the $(2, \pm 2)$-modes of the $+$-polarisation of the strain data from these waveforms, which are produced using `pycbc`, as described in section 7.1. The training data is produced in a similar fashion to the training data for the previous examples, substituting the `IMRPhenomP2` model for NR data. The NR waveforms are sampled at $1024\,\mathrm{Hz}$ and nominally evaluated at a total system mass of $60\,\mathrm{M}_{\odot}$. The strain values are then multiplied by a factor of 10^{19}, and the time values by 100, which were found to substantially reduce the time required to run the optimisation and matrix inversion steps during the training and evaluation of the model respectively. The values of the various spin components already lie within a range $[-1, 1]$, and so these were left unchanged. As in the previous examples the model was trained on the natural log of the mass ratio, which was again found to produce a better surrogate model.

The outputs of the model must be suitably rescaled to provide waveforms corresponding to the correct total system mass, and in physical units.

The outputs of the model which are presented in this chapter assume that the source is observed "face-on", and therefore there is no variation in the angle at which the source is observed.

The covariance structure of the Heron model follows the structure of the previous examples, but adds an extra four dimensions to account for the additional spin parameters in the NR data, giving a covariance function

$$K = \mathsf{C} \times \mathsf{SE}^{(t,\log q,s_{1x},s_{1y},s_{1z},s_{2x},s_{2y},s_{2z})} \tag{7.6}$$

This choice of covariance function follows the considerations made in the previous models: with none of the dimensions expected to be discontinuous the squared-exponential kernel is a reasonable choice for each.

7.4 | Verification of the model

The sparsity of training data poses a considerable challenge to the testing and verification of a model such as the Heron model; conventional approaches to testing such a model involve setting aside a fraction of the training data to compare to the model output when evaluated at the parameter space location of each test datum.

The quantity of numerical relativity waveforms available at present in the Georgia Tech catalogue makes this approach difficult, as some regions of the parameter space are very sparsely sampled, and omitting a training waveform in this location may significantly complicate the process of training the model. To overcome this I have carried out four separate categories of test on the Heron model.

In-sample tests where the entire catalogue of available training waveforms are used to condition the GP used by the model. Waveforms are then produced from the model at the parameter locations which correspond to each of the training waveforms, and the match between the Heron waveform and the NR waveform is calculated.

Out-of-sample tests where a single waveform from the catalogue is omitted from the set of training waveforms used to condition the GP, but the hyperparameters which were determined during training of the full model are used. A GP model is conditioned on a reduced catalogue for each waveform, and the waveform is produced from the reduced Heron model which corresponds to the omitted NR waveform. The match is then computed between these two waveforms.

Tests against phenomenological models where the match is computed between waveforms produced by Heron and by other waveform models, such as `SEOBNRv3` and `IMRPhenomPv2`.

Tests against other NR catalogues where the match is computed between waveforms from another NR catalogue and a waveform corresponding to the same parameters produced by the Heron model.

Each approach to testing has different advantages and disadvantages, and test for different aspects of the model's performance.

7.4.1 In-sample tests of the Heron model

The simplest set of tests which I perform on the Heron model are *in-sample* tests, which effectively test the model's ability to reproduce its own training data. For the Heron model this involved computing the mean waveform from the GP corresponding to each waveform which was used in the training set. The match was then calculated between each mean waveform and the corresponding NR training waveform using the expression for waveform match, $\mathcal{M}$, given in equation 7.3.

In-sample testing ought to reveal problems with the choice of hyperparameters in the model, or inconsistencies in the training data itself.

Figure 7.4.1 plots the histogram of the mismatch (equal to $1 - \mathcal{M}$) values which resulted from these tests against the Georgia Tech waveforms used as the training data (plotted as the black-outlined histogram). Reassuringly the mismatch between the vast majority of the model outputs and the training data are small; a small number of outliers appear to have a mismatch greater than 0.01. Figure 7.4.2 plots the Euclidean distance between each sampled point and the next-nearest training datum (calculated as the norm of the vector between the two points) and the mismatch. This reveals that a small number of points which are very close to other points in the training data are generally responsible for these high mismatches. This may indicate that the model is struggling to fit a smooth function to a small number of closely-spaced waveforms. Further work is likely to be needed to identify whether these outliers are the result of flawed training data, or a some other problem with the model.

The distribution of mismatches between the Heron model and the NR data shows that for the majority of waveforms in the test the mismatch is smaller than between the NR data and either of the approximant models. This is reassuring, however, since the approximant models are conditioned on NR data, and are not necessarily conditioned on waveforms at the same locations as the test waveforms (whereas the Heron model *is* for this in-sample test), this does not provide a comprehensive measure of the Heron model's effectiveness as a surrogate.

7.4.2 Out-of-sample tests of the Heron model

A more rigorous test of a predictive model involves comparing the model's output in a region of the parameter space which does not contain a training datum. This process, known as out-of-sample testing, is difficult for the Heron model, thanks to the large (seven dimensional) parameter space,

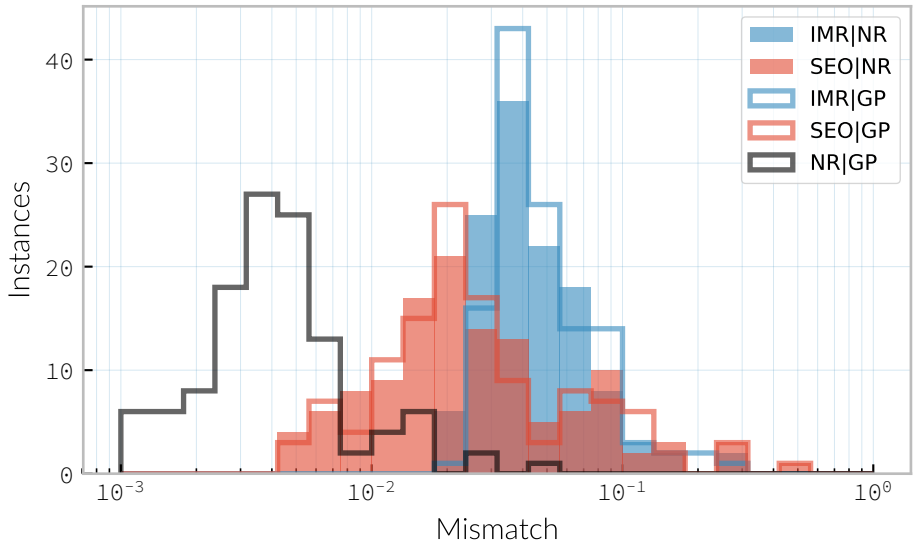


Figure 7.4.1: The distributions of mismatches between mean waveforms from the Heron model and each of the NR waveforms from the Georgia Tech waveform catalogue (black outline histogram) used in the training set using the procedure described in section 7.4.1. Additionally, the mismatch distributions between waveforms produced at the same parameters as the NR waveforms by the `SEOBNRv3` (red outline histogram), and the `IMRPhenomPv2` (blue outline histogram) phenomenological waveform models are plotted. For comparison the distributions of mismatch between the same Georgia Tech waveforms and the corresponding waveforms from the `SEOBNRv3` and `IMRPhenomPv2` models are plotted as solid red and blue histograms respectively.

and the small number of available training waveforms (132). As a result, removing a substantial fraction of the waveforms in order to produce a set of test data would be likely to substantially affect the predictive power of the model.

To overcome this I have performed a *leave-one-out* testing procedure. In order to do this multiple training datasets are produced; from each a single waveform is omitted. This reduced dataset is then substituted for the data on which the full Heron model's GP is conditioned, retaining the same hyperparameter values as the full model. The reduced Heron

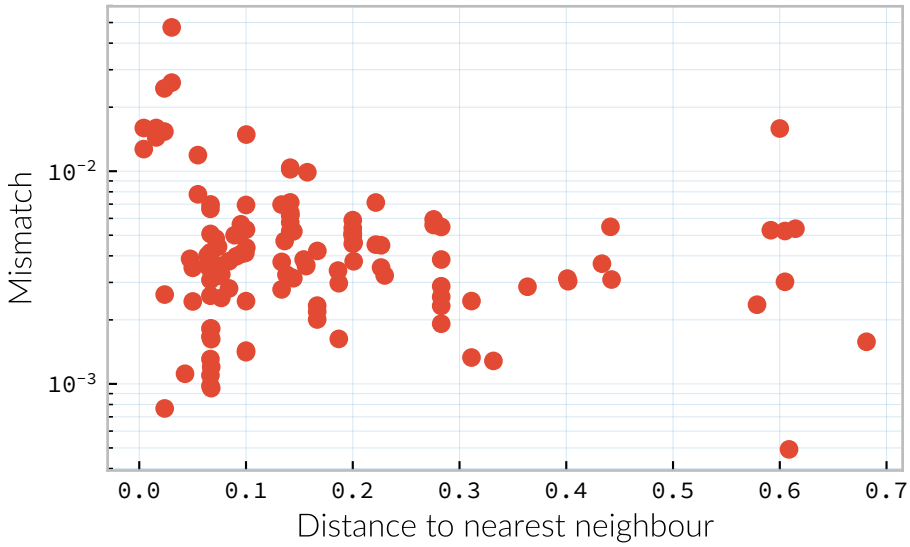


Figure 7.4.2: The mismatch between mean waveforms produced by the Heron model and NR waveforms compared to the distance to the next-nearest training waveform in the training data. Here the distance is calculated simply as the Euclidean distance between the two points in parameter space.

model is then evaluated at the parameter location corresponding to the omitted waveform, in order to compute a predicted mean waveform. The mismatch between the predicted waveform and the omitted NR waveform was then computed, and the distribution of these mismatches is plotted in figure 7.4.3 as a black-outlined histogram, with the mean mismatch from these tests being 0.035, with 95% of the tests producing a mismatch between 0.114 and 0.0001.

In contrast to the in-sample tests from section 7.4.1 the out-of-sample tests do give some indication of the model's capability as a surrogate model. The mismatches between the leave-one-out model and NR data in 7.4.3 show that the Heron model is capable of producing waveforms which have a small mismatch with the NR waveforms, with the vast majority having a mismatch smaller than 0.1. The maximum SNR, ρ_{eff}, at which the true

waveform and the model waveform cannot be distinguished by a parameter estimation process is a function of the mismatch $(1 - \mathcal{M})$ (that is, below an SNR of ρ_{eff} the parameter estimation process is not worsened by the use of the surrogate compared to an NR waveform), with ρ_{eff} defined [281, 282]

$$\rho_{\text{eff}} = \frac{1}{\sqrt{2(1 - \mathcal{M})}}. \tag{7.7}$$

Therefore the mean mismatch (0.035) from these tests corresponds to a $\rho_{\text{eff}} \approx 3.8$, with 95% of the tests corresponding to a range $\rho_{\text{eff}} = (2.09, 22.36)$, which is adequate for advanced LIGO parameter estimation, where SNRs will be low, but would be insufficient for detectors such as Laser Interferometer Space Antenna (LISA), which are expected to observe high SNR signals.

The variation of mismatch over the parameter space of the Georgia Tech catalogue is explored in figure 7.4.4, with waveforms with higher mismatch generally being in areas which are separated from the main body of training data. There is no clear pattern to locations with high waveform mismatch in the parameter space, although low-spin equal-mass systems can be seen to produce a low mismatch, while systems with non-aligned spins generally show higher mismatch than systems with either no spin or aligned-spin.

7.4.3 Tests against other models

It may also be helpful to understand how the outputs of the Heron model compare to conventional phenomenological approximants which are in widespread use. To do this I calculated the mismatch between the output of the Heron model at the same parameter locations as the in-sample and leave-one-out tests.

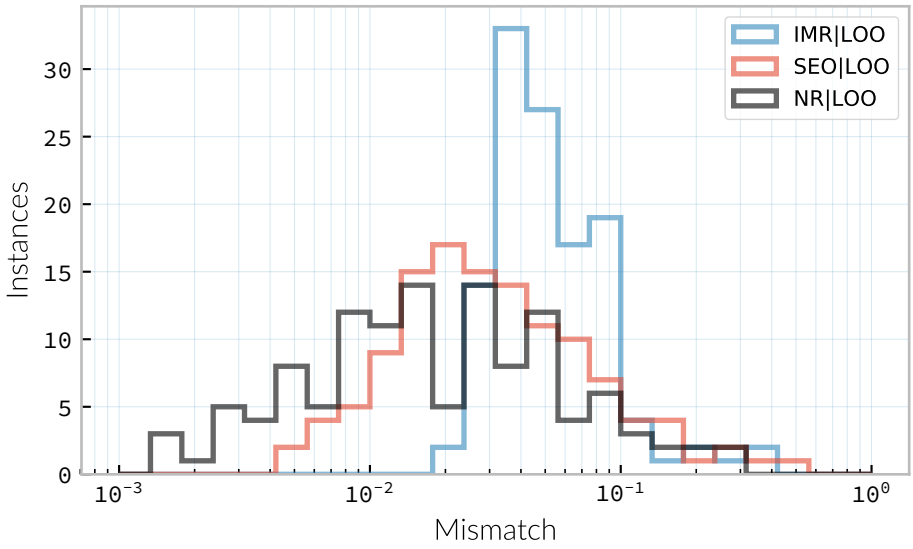


Figure 7.4.3: The distributions of mismatches between mean waveforms from the Heron model and each of the NR waveforms from the Georgia Tech waveform catalogue (black outline histogram) used in the training set using the leave-one-out (LOO) testing procedure detailed in section 7.4.2. Additionally, the mismatch distributions between waveforms produced at the same parameters as the NR waveforms by the `SEOBNRv3` (red outline histogram), and the `IMRPhenomPv2` (blue outline histogram) phenomenological waveform models are plotted.

The results of these tests are plotted in figures 7.4.1 and 7.4.3 for the in-sample and leave-one-out scenarios respectively. In addition, figure 7.4.1 shows the distribution of mismatches between the training NR data and each of these models (as solid histograms). In the in-sample case the Heron model reproduces the NR waveforms with substantially lower mismatch than either phenomenological model. This behaviour is to be expected, since the Heron model has direct access to the NR data, where the phenomenological models do not. It is worth noting that the mismatch for `SEOBNRv3` is consistently smaller than that of `IMRPhenomPv2` against both NR and the Heron model. `IMRPhenomPv2` is known to be accurate over a smaller range of black hole spins than the `SEOBNRv3`

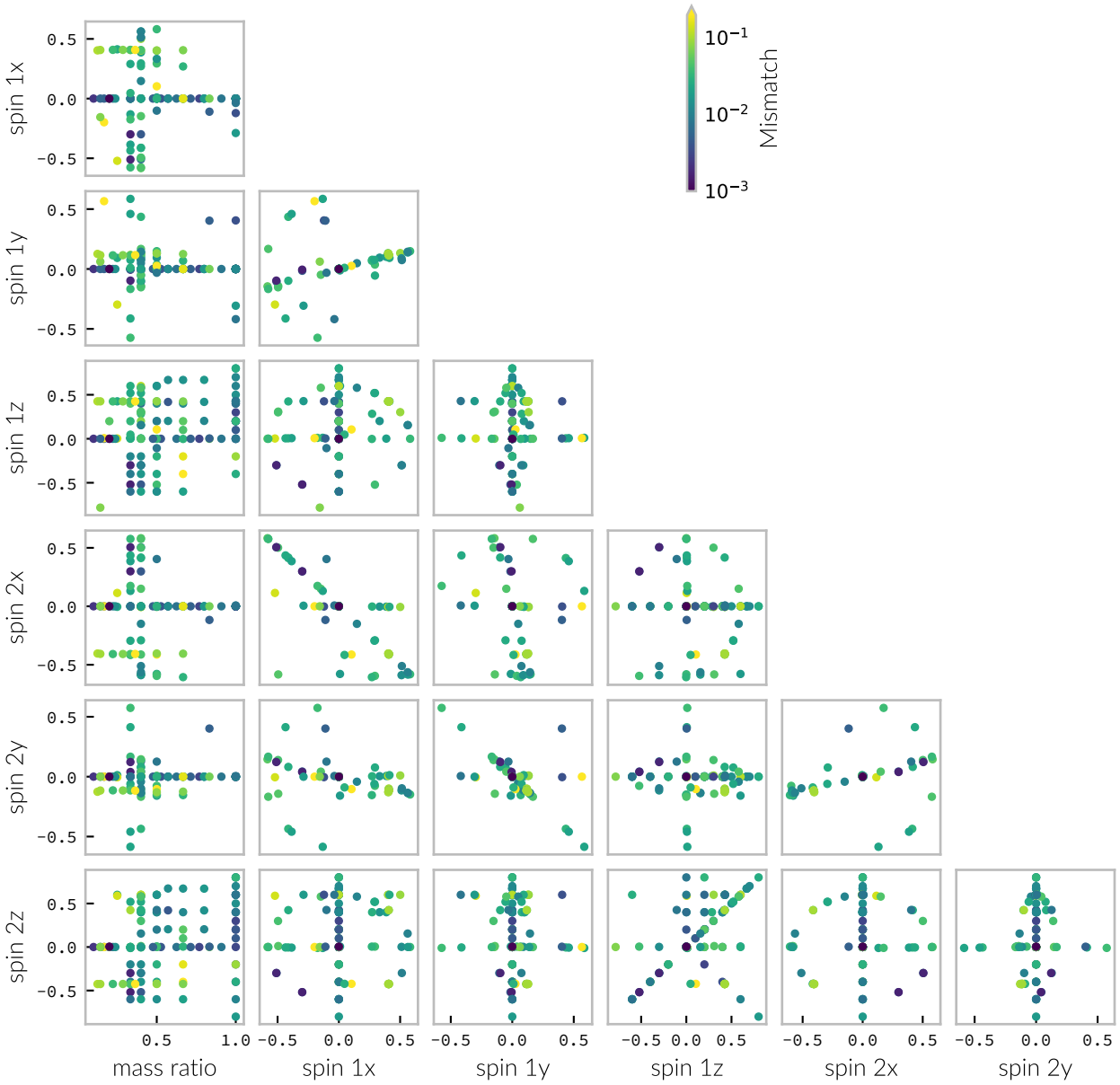


Figure 7.4.4: A corner plot across the BBH parameter space showing the leave-one-out mismatch, as described in section 7.4.2, between the mean waveform produced by each (reduced) Heron model, and the corresponding omitted Georgia Tech NR waveform.

model. The leave-one-out distributions are more interesting, as the Heron model here produces a distribution of mismatches which is comparable to SEOBNRv3, with tail of smaller mismatches, indicating that the model retains comparable predictive power to the phenomenological model.

7.4.4 Tests against other catalogues

While removing waveforms from the Georgia Tech catalogue to hold back for testing data may be detrimental to the Heron model, a number of other waveform catalogues are available, including the extensive SXS catalogue [158, 159, 160, 161, 162, 163] (and described in section 2.3.4.2).

Ideally an GP model would include waveforms from both catalogues, in order to account for systematic uncertainties between the two catalogues. However, given that the Heron model is trained only on the Georgia Tech catalogue it is possible, with care, to compare the output of the Heron model to NR waveforms in the SXS catalogue. This comparison must be interpreted with care, however, due to the presence of said systematic uncertainties, and the mismatch between the Heron model and the NR waveforms will not take this into account.

The matches between 243 waveforms from the SXS catalogue and the corresponding mean waveform from Heron are plotted in figure 7.4.5 showing the mismatch across the parameter space. The distribution of these mismatches is plotted in figure 7.4.6 (as a black-outline histogram) with the mismatch distributions of the two phenomenological models at the same parameters plotted for comparison. Again the Heron model produces a distribution comparable to SEOBNRv3. As with the distribution plotted in figure 7.4.4 there are no clear patterns in the instances of high-mismatches across the parameter space, although non-spinning waveforms generally produce the lowest mismatches.

The mean mismatch in these tests, between the waveforms produced by Heron and those from the SXS catalogue is 0.034, with 95% of the tests having a mismatch between 0.112 and 1.2×10^{-4}. This corresponds to an effective SNR limit (as defined in equation 7.7, representing the highest SNR at which the NR and Heron-derived waveforms are indistinguishable) between 2.11 and 63.7 across the parameter space tested.

7.5 | Example waveforms

While I have discussed at length the various tests which I carried out on the Heron model, it is valuable to be able to visually compare the output of this model with the phenomenological models used in testing.

Figures 7.5.1, 7.5.2, 7.5.3, and 7.5.4 illustrate the output of the model in a variety of scenarios, respectively for a non-spinning , non-equal mass system; for a non-spinning, equal mass system; a system with aligned spins; and a system with non-aligned spins. In each of these figures the mean waveform for a given parameter space location is plotted, alongside 50 individual draws from the GP. `IMRPhenomPv2` and `SEOBNRv3` waveforms are overlaid on draws from the model, alongside its mean prediction, and the variance of the prediction.

I have plotted the distribution of mismatches between the model predictions and the two phenomenological approximants are shown in the right panel of each figure, with matches calculated between the approximant waveforms (plotted as solid lines) and one-hundred sample waveforms drawn from the model (plotted as histograms).

Figure 7.5.1 shows a waveform produced from the various models at the same parameter location as one of the Georgia Tech NR waveforms, which is plotted alongside for comparison, and shows that the GP reproduces the NR waveform well with the mean waveform, however individual draws

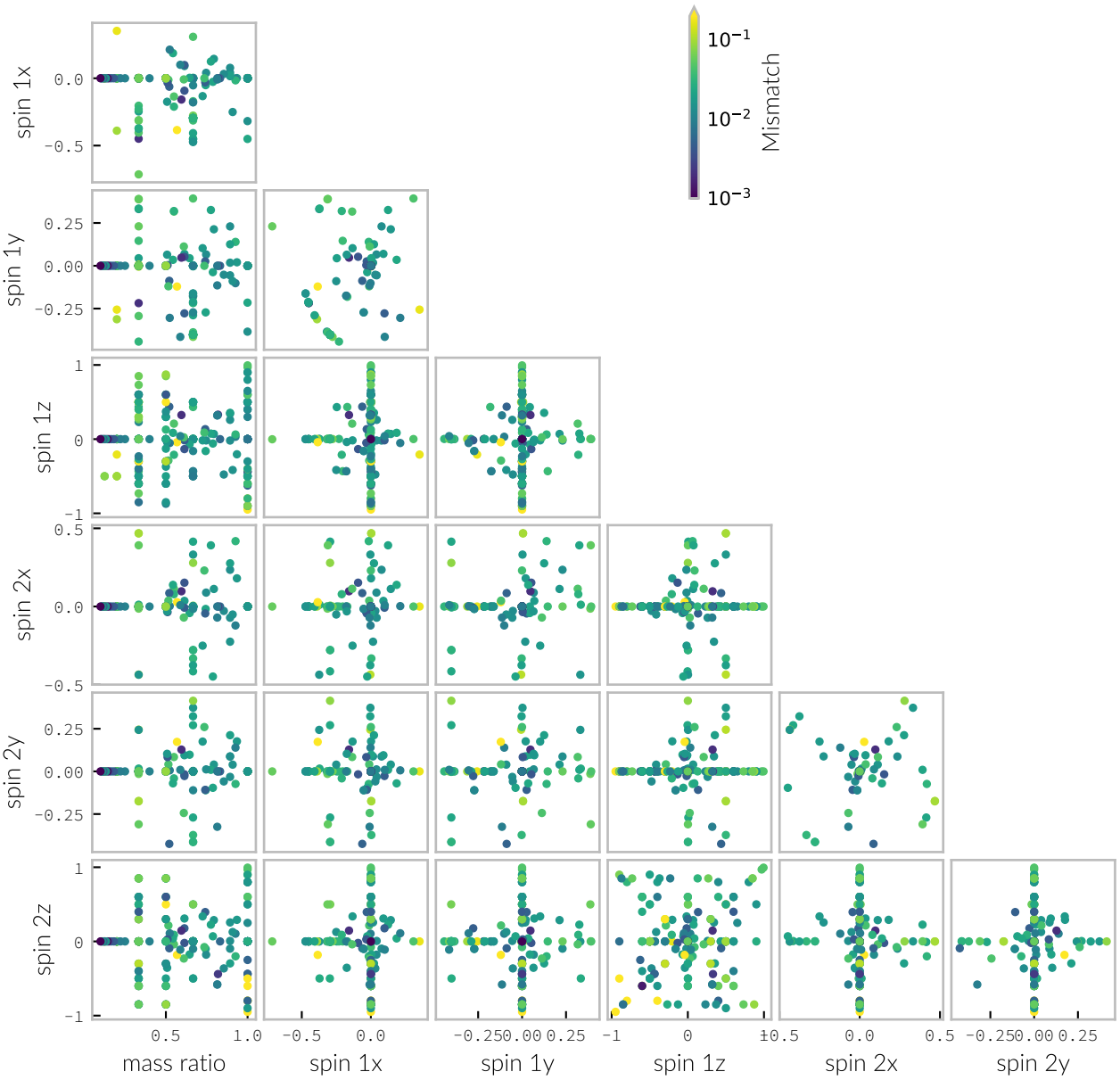


Figure 7.4.5: A corner plot across the parameter space of the SXS waveform catalogue showing the mismatch between the mean waveform produced by the Heron model and each SXS NR waveform, resulting from the testing procedure detailed in section 7.4.4.

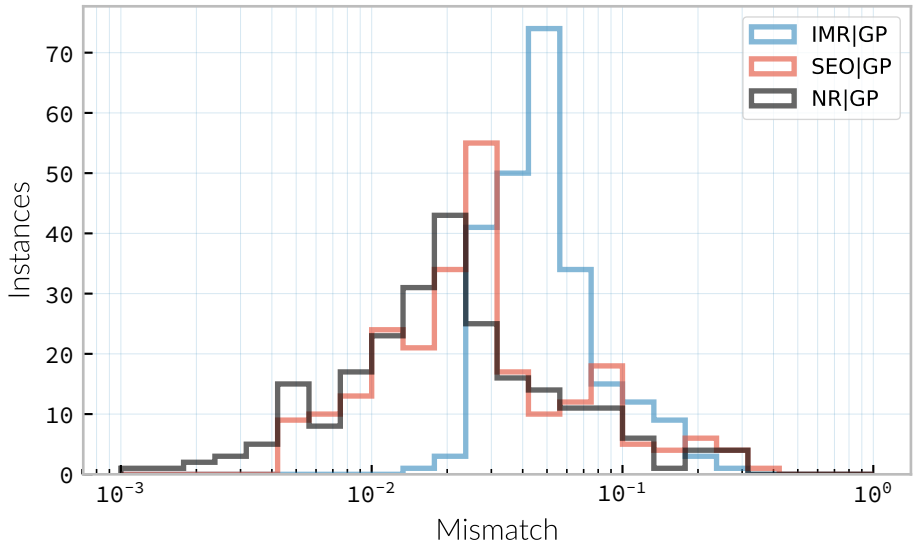


Figure 7.4.6: The distributions of mismatches between mean waveforms from the Heron model and each of the NR waveforms from the SXS waveform catalogue (black outline histogram) using the testing procedure detailed in section 7.4.4. Additionally, the mismatch distributions between waveforms produced at the same parameters as the NR waveforms by the `SEOBNRv3` (red outline histogram), and the `IMRPhenomPv2` (blue outline histogram) phenomenological waveform models are plotted.

deviate from the NR waveform. This behaviour can be seen more clearly in the histogram of mismatches between the individual draws and the two approximant waveforms, where the individual mismatches are invariably worse than the mismatch for the mean waveform (which are marked as solid lines of the appropriate colour for each approximant).

It is noteworthy that this mismatch is smaller than the mean of the mismatches between the sample draws and the phenomenological models. This is a result of the mismatch being a somewhat asymmetric indicator: the mismatch will always be higher for a waveform which over-estimates or under-estimates some feature of the waveform, where the over- and under-estimates will be averaged through the use of the mean waveform, producing a lower mismatch.

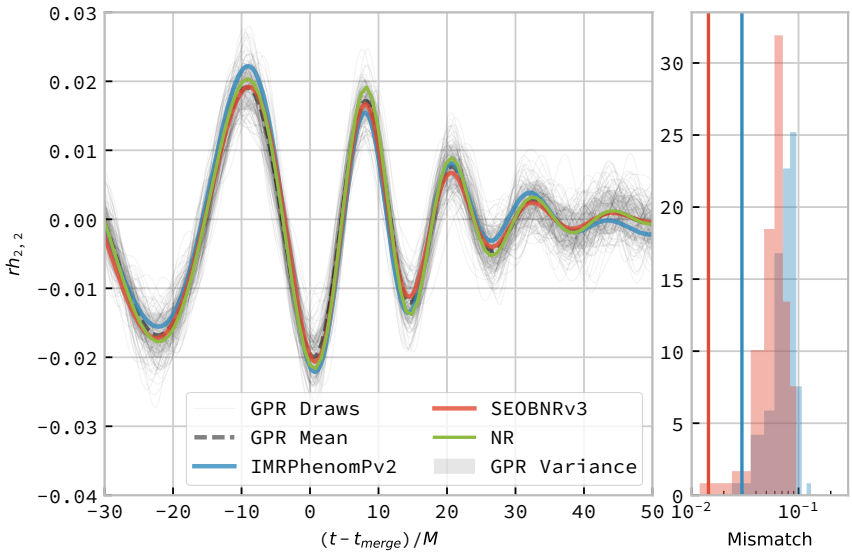


Figure 7.5.1: GP **regression predictions, compared to** NR. One hundred draws from the Gaussian process (left panel) for a non-spinning configuration ($\boldsymbol{s_1} = (0, 0, 0)$, $\boldsymbol{s_2} = (0, 0, 0)$, $q = 0.625$), shown as light grey lines compared to two analytical approximant models, `SEOBNRv3` and `IMRPhenomPv2` in red and blue respectively. The mean draw from the Gaussian process is shown as a grey dashed line, while the associated variance is plotted as a grey-filled region surrounding the mean. The differences between the phenomenological model and the GP regression model waveforms are seen to also exist between the phenomenological model waveforms and the NR-derived waveform (`GT0374`), plotted here in green. In the right panel the distribution of mismatches between the samples and both phenomenological waveforms are shown, with the vertical lines representing the mismatch between the GP regression and the phenomenological waveform.

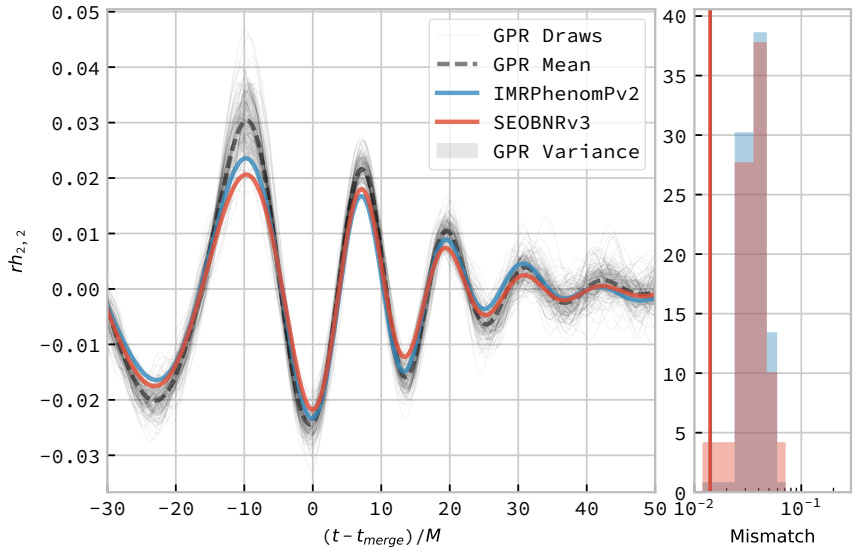


Figure 7.5.2: Non-spinning waveform. One hundred draws from the Gaussian process (left panel) for a non-spinning, equal-mass configuration ($\boldsymbol{s_1} = (0, 0, 0)$, $\boldsymbol{s_2} = (0, 0, 0)$, $q = 1.0$), shown as light grey lines compared to two analytical approximant models, `SEOBNRv3` and `IMRPhenomPv2` in red and blue respectively. The mean draw from the Gaussian process is shown as a grey dashed line, while the associated variance is plotted as a grey-filled region surrounding the mean. In the right panel the distribution of mismatches between the samples and both phenomenological waveforms are shown, with the vertical lines representing the mismatch between the GP regression and the phenomenological waveform.

While figures 7.5.1, and 7.5.2, which are both waveforms for non-spinning systems, produce low mismatches, and waveforms which are visually comparable to the two approximant models' waveforms, the waveform presented in 7.5.3, for an anti-aligned waveform, is both visually worse (with the mean Heron output overlapping with neither approximant) and the mismatches are larger than the previous cases. Here the uncertainty of the prediction is also larger, indicating the model's lower confidence in the predicted waveform, which is vindicated by the lack of agreement between both the mean waveform and either of the approximants.

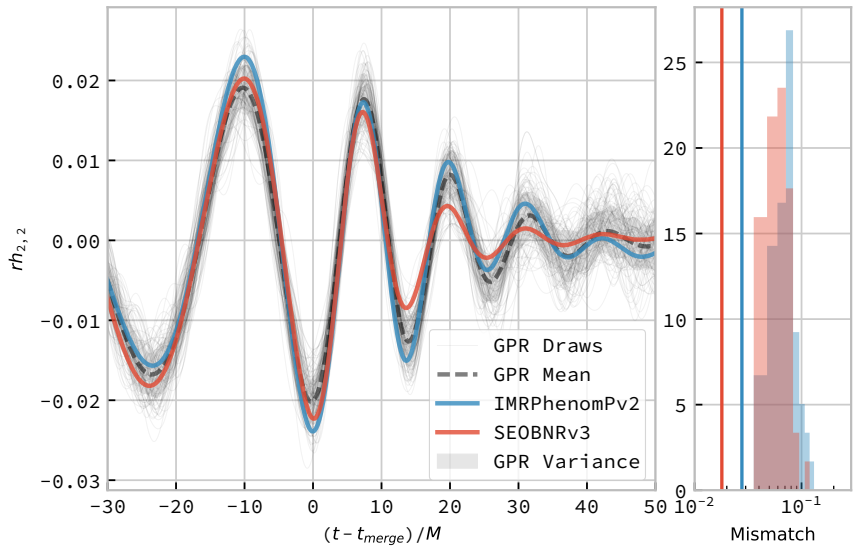


Figure 7.5.3: **Anti-aligned spin waveform.** One hundred draws from the Gaussian process (left panel) for a non-spinning, equal-mass configuration ($\boldsymbol{s_1} = (0, 0, 0.6)$, $\boldsymbol{s_2} = (0, 0, -0.6)$, $q = 1.0$), shown as light grey lines compared to two analytical approximant models, `SEOBNRv3` and `IMRPhenomPv2` in red and blue respectively. The mean draw from the Gaussian process is shown as a grey dashed line, while the associated variance is plotted as a grey-filled region surrounding the mean. In the right panel the distribution of mismatches between the samples and both phenomenological waveforms are shown, with the vertical lines representing the mismatch between the GP regression and the phenomenological waveform.

Figure 7.5.4, which represents a waveform for a precessing (face-on) system shows even greater uncertainty than 7.5.3, and an associated increase in both the mismatch between the approximants and the mean output, and between the approximants and the individual draws.

7.6 | Summary

We have entered the era of routine GW detection, and the ability to accurately and rapidly characterise signals from events such as coalescences

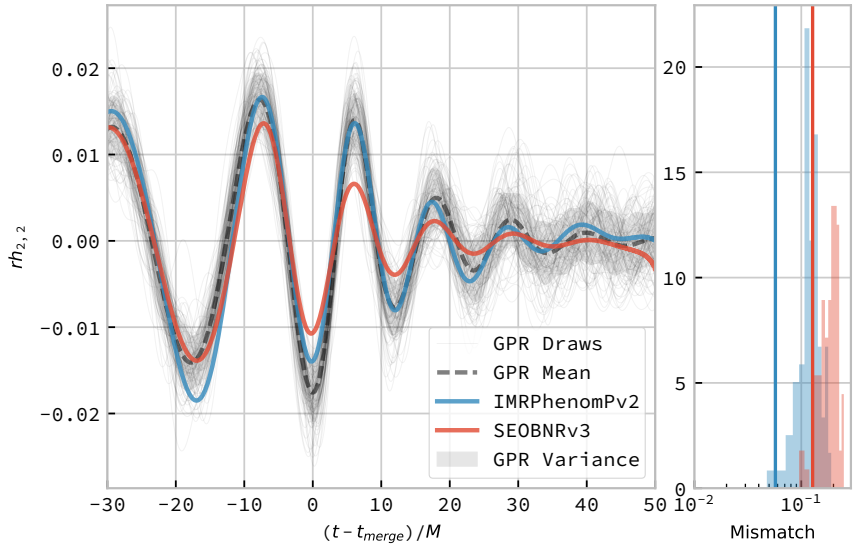


Figure 7.5.4: **Precessing waveform**. One hundred draws from the Gaussian process (left panel) for a precessing system, with a mass ratio $q = 0.25$, and a spin configuration ($\boldsymbol{s_1} = (0.35, 0.1, 0.2)$, $\boldsymbol{s_2} = (0, 0, 0.4)$), shown as light grey lines compared to a single analytical approximant model, `IMRPhenomPv2` in blue. The mean draw from the Gaussian process is shown as a grey dashed line, while the associated variance is plotted as a grey-filled region surrounding the mean. In the right panel the distribution of mismatches between the samples and both phenomenological waveforms are shown, with the vertical line representing the mismatch between the GP regression and the phenomenological waveform.

will be critical to understanding the properties of these systems. This characterisation process relies on the availability of waveform templates which are either pre-computed prior to the analysis being run, or can be generated on-the-fly. Highly accurate waveforms, generated by simulations, are able in principal to facilitate accurate inference on detected signals. However, the expense of producing them limits their coverage of the parameter space; as a result of this lack of coverage, and the considerable time requirements to produce new waveforms, any inference method which relied solely on techniques could not hope to satisfy the requirement to

rapidly characterise signals, and would not be practical in a scenario where multiple events are detected every month. Phenomenological models, which can be evaluated rapidly, are available, which attempt to interpolate across a large volume of the parameter space, but the accuracy of the waveforms which they produce can be difficult to assess. This leads to the possibility of introducing biases into the inferred properties of the system which generated the signal.

In this chapter I have laid-out an approach to producing a waveform model which is capable of producing estimates of the uncertainty in the waveform in addition to estimates of the gravitational waveform in regions of parameter space which have not been explored by NR simulations. The ability to include these uncertainties should allow the accuracy of gravitational wave parameter estimation to be improved, as it allows biases introduced by uncertain waveform production to be taken into account.

In contrast to previous attempts to produce a GP based model for waveforms, such as [283], this model is trained on data from the Georgia Tech waveform catalogue, described in section 7.1. Other approaches have also been taken to producing surrogate models – eg the NRSur family of models [174, 175, 176] (discussed in section 2.3.6) – however these approaches do not have the attractive feature of producing an estimate of the waveform uncertainty.

In section 7.4 I presented a number of tests which were used on the model to ensure that its outputs are reliable. The results of these tests indicate that the Heron model is capable of producing similar matches against numerical relativity waveforms

I have presented a number of waveforms which have been produced by the Heron model in section 7.4, and make comparisons between its output and two phenomenological models. These comparisons show a difference between the behaviour of the two models which is most pronounced

during the inspiral section of the waveform. This difference also occurs between the phenomenological model and the waveform produced from GP regression. A number of phenomena are likely to have contributed to this discrepancy. One such difference in the systematic errors of the simulations used to produce the training data for the model compared to those used to calibrate the phenomenological models. Additionally, the relatively small number of waveforms used to calibrate the phenomenological models compared to the model are likely to introduce systematic errors in the waveforms produced by those models. In order to reduce the effect of systematic errors from a larger model could include waveforms from a number of different waveform catalogues, however the addition of more waveforms will increase the memory requirements to both train and evaluate the model. The waveform model tends towards producing conservative estimates of the waveform, this is clearly visible in the variance of the precessing waveform in figure 7.5.4. The use of additional waveforms is likely to improve the confidence of the model's prediction.

In order for a GP based approach such as this to be practical for parameter estimation studies using data from LIGO or Virgo it would be necessary to have a means of producing waveforms which are capable of modelling a greater amount of the inspiral than the Heron model can currently provide. One potential approach to solving this problem is hybridising the output waveform from the model with waveforms produced from a post-Newtonian approximant, in a similar manner to that used by [159].

The test which are used on the Heron model, and presented in this chapter, use the time and phase optimised match. While this is a metric which is frequently used to understand the behaviour of waveform models, it is not completely thorough, and in the future it would be desirable to develop additional tests which are more sensitive to differences in amplitude

between the reference and model waveform, for example. Additional attention is also required for the outliers in the in-sample tests, in order to better understand whether these are the result of problems originating in the model, for example, through an incorrect choice of hyperparameter values, or are the result of problems with the training data itself.

The approach which was taken to produce the Heron model, in which time-domain waveforms were used directly as training data, has the major advantage that the posterior GP distribution is Gaussian, which allows uncertainties from waveforms to be handled in a fairly straight-forward manner. This would be especially helpful if the Heron model were used as part of the likelihood calculation during a parameter estimation process. The major disadvantage which comes with this approach is the complexity of the waveform in the time domain, compared to producing models over the amplitude and phase of the waveform separately. In this case the uncertainty will not be Gaussian, and in order to incorporate waveform uncertainty individual waveform draws would be required in the parameter estimation process. This will be significantly slower, requiring many more evaluations of the dot product between waveform draws and the data than would be required between the mean waveform and the data. It can also be seen from the examples in section 7.5 that the individual draws will produce a different mean mismatch than the mean waveform will. For a model which used this approach to be useful for parameter estimation studies more work would be required to properly understand and characterise the behaviour of the individual waveform draws.

All of the waveforms which are presented in section 7.5 are "face-on" waveforms. In order to allow variation of the source orientation both h_+ and $h_\times$ waveforms are required. The most straight-forward approach to producing both of these waveforms is likely to be to produce two GP surrogate models, one for each polarisation, and then combine these to

produce an appropriate waveform.

While the prototype version of the Heron model which is presented in this chapter is clearly capable of producing waveforms, additional development is still required for this to be a model which is useful for parameter estimation studies.

8 Summary and conclusions

In this thesis I have attempted to provide an overview of the broad field of gravitational wave (GW) research, ranging from the development of detectors capable of making observations of the minute displacements produced by propagating gravitational waves, to the various astrophysical systems which are expected to produce observable GWs.

The arrival of the first GW detection in September 2015 ushered-in a new era of observational GW astronomy, and with it a number of new challenges. The work presented in this thesis is an attempt to address some of these challenges. The development of the Heron waveform model, which is detailed in chapter 7, is an attempt to address the need both for more accurate models of binary black hole (BBH) signals, but also to incorporate understanding of the uncertainty of these models into analyses which make use of them. Heron, as a model which is fundamentally probabilistic in nature, provides a method for interpolating between highly-accurate numerical relativity (NR) waveforms, in order to give model coverage over a large region of the intrinsic parameter space for BBH systems, while also providing a measure of the uncertainty of the interpolant, which can then be included in the overall uncertainty associated with the analysis. In this thesis I demonstrate the ability of this model to produce BBH waveforms across their parameter space, and make comparisons between the output of the Heron model and other approximant models. Much work remains in order to make this waveform

suitable for use in parameter estimation studies, or other analyses. While the accuracy of the model is currently tested by calculating mismatch between its outputs and other models' outputs there is scope for additional testing to further validate the model, and ensuring that the model produces parameter estimates for known signals which are consistent with those found using other approximant models would be essential for Heron to be an alternative to these. There is much potential for future developments in the model, including properly handling source orientation, producing waveforms with longer inspirals, and incorporating data from more than one NR catalogue in an attempt to account for systematic biases which may be present in an individual catalogue, as well as providing a larger quantity of training data.

The first multimessenger observation involving GWs occurred in August 2017, with the observation of GW170817, and its associated short gamma ray burst (sGRB), and later emission across the electromagnetic (EM) spectrum. Multimessenger observations provide new possibilities for inferring properties and understanding the mechanisms of highly energetic astrophysical systems. One example of such a system is the production of sGRBs. GWs can provide insight into the progenitors of these events, such as binary neutron star (BNS) events. In chapter 5 I discuss the development of a technique which is capable of inferring the opening angle of sGRBs based on the observed rate of both BNS and sGRB events in the local universe. While this technique, which was developed before the observation of GW170817 and its associated sGRB, makes some assumptions which may seem naive in light of the unusual nature of this sGRB (specifically my choice of a top-hat jet model, and the low luminosity of GRB 170817A which has led to this model being disfavoured since 2017) its hierarchical nature will allow the technique to be adapted to other jet models, and for the model to be incorporated into a larger model which considers both a

larger quantity of data provided by GWs detectors, and observations from other regions of the EM spectrum.

The approach outlined in chapter 5 relies on understanding the rate of GW events; in order to infer this rate the sensitivity of the GW detectors must be well-understood. This task is non-trivial, as the noise present in measurements from these detectors is both non-stationary and evolves over time. The range of the detectors, which is affected by their sensitivity, is measured by "injecting" known signals, produced from theoretical models, into real, noisy data from the detectors. By injecting signals at a variety of amplitudes, corresponding to a range of distances, and determining which can be recovered by an analysis, allows the sensitivity of the detector over time to be estimated. A technique for producing these "mock data challenges" (and its implementation in the Minke python package) for transient signals is described in chapter 3, along with an overview of how the unmodelled searches, which these challenges are intended to test, operate.

The field of GW astronomy is developing rapidly, but as a young field, the future is likely to bring many surprises, and there are likely to be many sources of GWs which have yet to be considered. In chapter 2 I discuss a number of sources which are considered likely to produce GWs, but in section 2.4.8.1 I present a preliminary study on the detectability of a source which has not been widely considered: encounters between black holes, and the potential capture of one black hole by another which results in a BBH coalescence event. I show that this source is likely to be detectable with the current generation of detectors in galaxies beyond the Local Group, and is a promising event type to be observed by burst search algorithms of the type discussed in chapter 3.

There is scope for the continuation of all of the work which was started for this thesis, from the development of the hierarchical sGRB beaming

angle analysis into part of a modern analysis for more sophisticated models, to the creation of a full waveform model from the Heron model. The results on the detectability of black hole encounters and captures makes it clear that a search for these signals may be a fruitful use of resources, while the use of Minke to determine the sensitivity of burst searches will continue throughout the advanced detector era, and potentially beyond.

A Gravpy

It is often useful to be able to produce estimate of the detectability of a given GW morphology in a GW detector. The robust and rigorous method for calculating whether a signal can be identified by a given detector and an associated search algorithm involves producing an mock data challenge (MDC), where copies of the signal are added into noise from the detector, or noise simulated from that detector's amplitude spectral density (ASD). The search algorithm is then run on these "injected" signals, and the efficiency with which it is capable of recovering the signals is used to determine the maximum distance at which a source may be detected.

This process is complicated, and normally time (and resource) consuming, and it can be helpful to be able to produce a rougher estimate of the detectable distance of a source, even if this is only to determine if a more detailed MDC study is worth-while.

In this appendix I will briefly introduce *gravpy*, a Python package which is capable of performing this task for arbitrary GW signals, and arbitrary GW detectors. Full documentation for the package is available online[1] which includes full installation instructions.

A.1 | Defining a detector

Normally we will wish to work with a pre-existing GW detector when performing an analysis. A number of detectors are implemented already

[1] At `https://code.daniel-williams.co.uk/gravpy`.

in the package, and can be accessed easily. For example

```
from gravpy import interferometers as ifo
aligo = ifo.AdvancedLIGO()
```

will load the ASD of the advanced Laser interferometer gravitational-wave observatory (LIGO) detectors at their design sensitivity.

Gravpy allows the ASD of a detector to be plotted easily. Running

```
import matplotlib.pyplot as plt
f, ax = plt.subplots(1,1)
aligo.plot(ax)
```

will produce a `matplotlib` plot of the ASD for advanced LIGO. (This function was used to produce a number of the ASD plots in chapter 1).

Additional detectors are available in the package (and are listed in the online documentation), however new detectors can easily be implemented by extending the package's `Detector` class, thanks to Python's object oriented design.

A.2 | Defining a source

In order to determine the detectability of a source we need to load its waveform. All of gravpy's supported waveforms are located in the `gravpy.sources` module, which can be imported using

```
import gravpy.sources as sources
```

it is then straight-forward to use any of the pre-implemented sources. For example, to simulate a BBH event such as GW150914:

```
from astropy import units as u
cbc = sources.CBC(frequencies=np.logspace(-4, 5, 1000) * u.hertz,
```

```
    m1=32*u.solMass, m2=30*u.solMass, r=0.8*1e9*u.parsec)
```

In much the same way as for detectors, this source can be plotted with

```
import matplotlib.pyplot as plt
f, ax = plt.subplots(1,1)
cbc.plot(ax)
```

While a number of (simple) sources are implemented in gravpy (such as BBH events with circular orbits and no spin) it is easy to add a new waveform by extending the `Source` class.

A.3 | Estimating an SNR

The signal-to-noise ratio (SNR) of a given source in a given detector can be easily worked out using the signal's `snr()` method. For example

```
print(cbc.snr(aligo))
```

reveals a rough estimate that GW150914 would have had an SNR around 112 in the advanced LIGO detector at its design sensitivity. This result should clearly be taken with caution, since the actual ASD of the design sensitivity is unknown, the BBH model used here was very simple, and assumed that the source was optimally oriented, and located in an optimal sky location, none of which were, in fact, the case for GW150914.

Performing the same operation on a variety of past, present, and future detector designs gives the SNRs in table A.3.1.

IFO	SNR
aLIGO	112.363
aLIGO [O1]	24.813
eLISA	109.124
Initial LIGO	6.379
GEO600	4.800
TAMA	0.258

Table A.3.1: The SNR of GW150914 estimated by gravpy had it been detected by a number of past, present, and future detectors. As noted in section A.3 these SNR calculations assuming optimal source orientation and sky location for each detector, and often use idealised or estimated values for the detectors' ASD. For comparison, GW150914 was detected in advanced LIGO with an SNR of 13 and 20 (for each detector) [58]. This event would not have been detectable in Initial LIGO, where an SNR of 8 would be required for a detection, agreeing with the findings of [58].

Glossary

Advanced LIGO The Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) is a second-generation interferometric gravitational wave observatory which is presently composed of two detectors located in the United States of America: LIGO Livingston Observatory, located in Livingston parish, Louisiana, and LIGO Hanford Observatory, located on the Department of Energy's Hanford Reservation in Washington. Both detectors are contained within the structures constructed for the initial LIGO instruments, and use the LIGO vacuum system (with some additions near the vertex). 3, 15

ALLEGRO A cryogenic resonant bar gravitational wave detector which was located at the State University of Louisiana, and operated until 2008. The detector consisted of a cryogenic bar of aluminium weighing 2500 kg with a resonant frequency around 904 Hz. ALLEGRO conducted joint GW searches with the initial LIGO network during the fourth science run [284]. 12

ALTAIR A cryogenic resonant gravitational wave detector which was located at the *Istituto di Fisica dello Spazio Interplanetario of Consiglio Nazionale delle Ricerche* in Frscati, Italy. The detector consisted of a 389 kg aluminium bar with a resonant frequency of 1763 Hz. 12

AURIGA A cryogenic resonant bar gravitational wave detector which was located in Padua, Italy, at the *Laboratori Nazionali di Legnaro*. The detector consisted of a 2300 kg aluminium bar held at 100 mK,

with a resonant frequency around 920 Hz when cooled. 18, 83

chirp mass ($\mathcal{M}$) A parameter of a compact binary system which determines the frequency evolution of the gravitational waveform emitted during the inspiral. It is defined as

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}. \tag{A.1}$$

45, 58

Cosmic Explorer A proposed third-generation interferometric gravitational wave detector [88]. 16, 27, 28

duty cycle (γ) The fraction of time during which a gravitational wave detector is actively observing, and is in a state suitable for producing analysable data. One of the major goals of detector development and characterisation is to increase the duty cycle of both individual detectors, and the overall duty cycle of the detector network.

135

effective spin (χ_{eff}) The effective spin is a dimensionless quantity which characterises the amount of spin in a compact binary system relative to the orbital angular momentum of the whole system. For two objects with masses m_1 and m_2, with spins $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ It is defined

$$\chi_{\text{eff}} = \frac{(m_1 \boldsymbol{s}_1 + m_2 \boldsymbol{s}_2) \cdot \hat{L}}{(m_1 + m_2)} \tag{A.2}$$

for $\hat{L}$ the unit vector of the orbital angular momentum of the system. 45

Einstein Telescope A proposed third-generation interferometric gravitational wave detector. 25, 27, 28, 78

EXPLORER A cryogenic resonant gravitational wave detector. 13

Fabry-Perot cavity An optical cavity, in which mirrors reflect electromagnetic radiation to form standing wave resonance. A Fabry-Perot cavity specifically has plane-parallel mirrors at either end of the cavity. 26, 30

Fermi A Gamma-ray space telescope operated by NASA and the US Department of Energy. Fermi is designed to carry-out all-sky surveys of gamma ray radiation using its wide-field Large Area Telescope (LAT) in addition to studying gamma ray burst (GRB) signals with the Fermi Gamma-ray Burst Monitor (GBM). 123, 124

Gaussian process A Gaussian process is a stochastic process in which every subset of the stochastic variables has a multivariate normal distribution. They are a flexible Bayesian regression method which are well suited to timeseries modelling, and as surrogates to poorly sampled functions. 157

GEO600 A German-UK interferometric gravitational wave detector located near Hannover in Germany. With shorter arms than other modern interferometric dectors (600 m) the detector has a lower sensitivity than its kilometre-scale brethren, but it has served as a testbed for the development of technology for these detectors. 13, 14, 26, 41, 52, 83, 229

GEO-HF The high-frequency upgrade of the GEO600 detector. 15

glitch A glitch is a transient noise event which is observed within a gravitational wave detector. These noise events are especially difficult to deal with, as they often resemble the burst-like signals which are searched for by some of the search algorithms operating on detector data. For a more detailed discussion see section 1.4.2. 38, 39

gravitational wave A propagating space-time metric perturbation. 230

GW150914 The first detected gravitational wave, produced by a BBH coalescence [105]. The detection was made on 14 September 2015 by the two advanced LIGO observatories during an engineering run, but while the detector was producing observation-quality data. GW150914 was announced to the general public on 11 February 2016. 3, 4, 42, 43, 89, 131, 224, 225, 230

GW151012 Formerly known as LVT151012, the second gravitational wave detection from a BBH system. At the time this event was detected (12 October 2015) it did not satisfy the criteria to be announced in the literature as a gravitational wave event, and so was denoted a "LIGO/Virgo Trigger". The event was widely known as the "Second Monday" event during the first observing run (GW150914 being the first Monday event). 3, 43, 232

GW151226 The second detected gravitational wave, produced by a BBH coalescence, and detected by the advanced LIGO network on 26 December 2015 [107]. Also known as the "Boxing Day Event" thanks to its detection late on Boxing Day. 3, 43, 131

GW170104 The first gravitational wave detection, made on 4 January 2017, during the second observing run of advanced LIGO, of a binary black hole coalescence [285]. 3

GW170608 A BBH event observed on 8 June 2017, during the second observing run of advanced LIGO [286]. 3

GW170814 The first gravitational wave event to be observed (on 14 August 2017) by a three-detector network, comprised the two advanced LIGO and the advanced Virgo detectors [287]. As a result it was (briefly) the best-localised gravitational wave event. 3, 44, 234

GW170817 The first GW event observed (on 17 August 2017) from a BNS event [84]. Also the first GW event to be observed alongside electromagnetic emssion, first gamma rays, then across almost the entire EM spectrum. 4, 24, 44, 125, 130, 150, 220

horizon distance ($\mathcal{D}_{\text{hor}}$) The greatest distance at which a GW source can be located in order to be confidently detected by a gravitational wave detector. The horizon distance is normally assumed to be the maximum detectable distance for an optimally oriented source in the region of sky which the detector is most sensitive to. It is then the distance which would cause a signal to be detected with an SNR of 12 in a detector, which is the required SNR for an unambiguous detection. The average sensitivity over the whole sky (since GW detectors' sensitivity varies over the sky thanks to their antenna pattern) and over all orientations of the source would yield a distance around 2.26-times smaller. The most frequently quoted horizon distance is the BNS horizon distance, which is a fairly unambiguous measure of a detector's sensitivity thanks to small anticipated variance in the masses (and hence signal morphologies) of these systems.

46, 132, 233

IMRPhenom A family of binary black hole phenomenological approximant models which are capable of producing waveforms for precessing systems. 64, 67, 68

inspiral The period of a binary orbit during which gravitational wave emission is produced as a result of orbital decay. 59, 63

Kagra Kagra is a 3-km advanced-generation interferometric detector which is under development at the Kamioka Observatory, Gifa Prefecture, Japan. It will be the first detector to be located underground,

in an effort to reduce the effect of Newtonian noise on the detector, and the first detector to contain cryogenic components, to reduce thermal noise from the optical suspensions and coatings. The present development timeline of Kagra will see it join the international detector network around the time that the advanced LIGO and Virgo detectors reach their design sensitivity.

15, 17, 25, 27, 46

LAGOS The Laser Antenna for Gravitational radiation Observation in Space was an early design study for a space-based GW detector, conceived in the 1980s. 16

LALSuite The LIGO Scientific Collaborations (LSC) Algorithms Library Suite is a large collection of C libraries and algorithms designed for a wide variety of GW analysis tasks. It is available to download from the LSC at `https://wiki.ligo.org/Computing/LALSuite`. 91

LVT151012 See GW151012. 3

Mario Schenberg A cryogenic (20 mK) resonant GW detector in Sao Paulo, Brazil. In contrast to most other resonant GW detector designs (but similarly to MiniGRAIL), it has a spherical shape in order to increase its sensitivity over the sky. 18, 232

MiniGRAIL A cryogenic (5 K) resonant mass GW detector located in Leiden, Netherlands. Smiliarly to Mario Schenberg this detector is spherical rather than the more conventional cylindrical shape of other resonant detectors.

18, 232

Minke A Python software package for designing and producing mock data challenges and datasets of simulated GW signals. Mock data chal-

lenges are essential for calculating the range (see horizon distance) of search algorithms which are run on detector data.

90, 91, 93

NANOGrav The North American Nanohertz Observatory for Gravitational waves is a scientific collaboration which aims to use pulsar timing arrays to detect very low frequency GWs. NANOGrav is based mainly in North America. 17

NAUTILUS A cryogenic resonant gravitational wave detector. 12, 18

null stream The null stream energy is the minimum amount of energy in whitened detector data which is inconsistent with a gravitational wave signal from a given sky location, across a network of detectors. 83, 85

Ricci scalar (R) A quantity representing the deviation in the area of an $(N-1)$-dimensional sphere embedded in a curved N-dimensional space, compared to the same sphere embedded in a flat N-dimensional space. 5

Ricci tensor A quantity which describes how the distance between two points within a volume varies as the volume is parallel-transported over a curved manifold compared to the same movement over a flat manifold. 5

Riemann tensor The tensor which describes the total curvature of a (Riemannian) manifold. At each point in the manifold it assigns a tensor which describes how much the metric locally at that point differs from the Euclidean metric. The Riemann tensor is essential to general relativity (GR), as it is used to describe geodesic deviation, the curvature of a "straight trajectory" in the presence of a gravitational field. 5, 17, 188

search pipeline A data analysis process, which may contain multiple separate components, which is capable of identifying candidate gravitational wave events ("triggers"), and performing inference on the signal, which may include determining the significance of the trigger, and parameter estimation. 93, 94, 96, 97, 130, 234

SEOBNR A family of binary black hole approximant models which employ the effective one body paradigm for producing waveforms, which are then calibrated against NR waveforms. 64

TAMA A first-generation interferometric gravitational wave detector which was located in Japan. 13, 83

TianQin A space-based GW observatory proposed by Sun Yat-sen University, which would consist of three spacecraft in geostationary orbit around the earth (in contrast to Laser Interferometer Space Antenna (LISA) which would be located at an Earth-Sun Lagrange point). 28

trigger An event candidate which has been identified by a search pipeline. 83, 85, 130, 131

Virgo An advanced-generation interferometric gravitational wave detector located in Cascina, Italy. While Virgo applies a similar Michelson-interferometer design to the advanced LIGO detectors, a number of design features of this detector are distinct, including the way in which optics are suspended. Additionally, the detector has a shorter arm-length: 3-km compared to the advanced LIGO 4-km arms. As a result the detector has a different noise profile from the LIGO detectors, and provided a valuable corroboration of the results from LIGO when it observed GW170814 independently. 3, 14, 24–27, 37, 41, 43, 44, 46, 52, 57, 74, 96, 136, 150, 230, 232

Bibliography

[1] D. Williams et al. “Constraints on Short, Hard Gamma-Ray Burst Beaming Angles from Gravitational Wave Observations”. In: ApJ 858 (May 2018), p. 79 (cit. on pp. xvi, 120, 121).

[2] Daniel Williams et al. “A Precessing Numerical Relativity Waveform Surrogate Model for Binary Black Holes: A Gaussian Process Regression Approach”. In: *arXiv e-prints* (Mar. 2019), arXiv:1903.09204 (cit. on pp. xvi, 69, 185, 190).

[3] Lev Davidovich Landau and E. M. Lifshitz. *The classical theory of fields*. 1975 (cit. on p. xxi).

[4] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. 2017 (cit. on pp. xxi, 6).

[5] B. P. Abbott et al. “GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs”. In: *Physical Review X* 9.3 (July 2019), p. 031040 (cit. on pp. 3, 43, 45, 131).

[6] R. A. Hulse and J. H. Taylor. “Discovery of a pulsar in a binary system”. In: ApJ 195 (Jan. 1975), pp. L51–L53 (cit. on pp. 4, 56).

[7] J. M. Weisberg and J. H. Taylor. “The Relativistic Binary Pulsar B1913+16: Thirty Years of Observations and Analysis”. In: *Binary Radio Pulsars*. Ed. by F. A. Rasio and I. H. Stairs. Vol. 328. Astronomical Society of the Pacific Conference Series. July 2005, p. 25 (cit. on pp. 4, 56).

[8] A. Einstein. “Die Feldgleichungen der Gravitation”. In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.* (1915) (cit. on p. 4).

[9] A. Einstein and N. Rosen. “On Gravitational Waves”. In: *Journal of The Franklin Institute* 223 (Jan. 1937), pp. 43–54 (cit. on p. 4).

[10] A. Einstein. “Über Gravitationswellen”. In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 154-167.* (1918) (cit. on p. 4).

[11] H. Weyl. *Raum, Zeit, Materie*. 1970 (cit. on p. 4).

[12] A. S. Eddington. “The Propagation of Gravitational Waves”. In: *Proceedings of the Royal Society of London Series A* 102 (Dec. 1922), pp. 268–282 (cit. on p. 5).

[13] F. A. E. Pirani. “On the Physical significance of the Riemann tensor”. In: *Acta Physica Polonica* 15 (Jan. 1956), pp. 389–405 (cit. on p. 5).

[14] John Archibald Wheeler and Kenneth Ford. *Geons, black holes and quantum foam : a life in physics*. 1998 (cit. on p. 6).

[15] C J Moore, R H Cole, and C P L Berry. “Gravitational-wave sensitivity curves”. In: *Classical and Quantum Gravity* 32.1 (2015), p. 015014 (cit. on pp. 11, 193).

[16] Jorge Cervantes-Cota, Salvador Galindo-Uribarri, and George Smoot. “A Brief History of Gravitational Waves”. In: *Universe* 2.3 (Sept. 2016), p. 22 (cit. on p. 12).

[17] H. Bondi. “Plane Gravitational Waves in General Relativity”. In: Nature 179.4569 (May 1957), pp. 1072–1073 (cit. on p. 12).

[18] J. Weber. “Detection and Generation of Gravitational Waves”. In: *Phys. Rev.* 117 (Jan. 1960), pp. 306–313 (cit. on pp. 12, 17).

[19] J. Weber. “Evidence for Discovery of Gravitational Radiation”. In: *Physical Review Letters* 22 (June 1969), pp. 1320–1324 (cit. on p. 12).

[20] J. Weber. “Gravitational Radiation Experiments”. In: *Physical Review Letters* 24 (Feb. 1970), pp. 276–279 (cit. on p. 12).

[21] J. Weber. “Anisotropy and Polarization in the Gravitational-Radiation Experiments”. In: *Physical Review Letters* 25 (July 1970), pp. 180–184 (cit. on p. 12).

[22] R. W. P. Drever et al. “Search for Short Bursts of Gravitational Radiation”. In: Nature 246 (Dec. 1973), pp. 340–344 (cit. on p. 12).

[23] J. L. Levine and R. L. Garwin. “Absence of Gravity-Wave Signals in a Bar at 1695 Hz”. In: *Physical Review Letters* 31 (July 1973), pp. 173–176 (cit. on p. 12).

[24] R. L. Garwin and J. L. Levine. “Single Gravity-Wave Detector Results Contrasted with Previous Coincidence Detections”. In: *Physical Review Letters* 31 (July 1973), pp. 176–180 (cit. on p. 12).

[25] J. L. Levine and R. L. Garwin. “New Negative Result for Gravitational Wave Detection, and Comparison with Reported Detection”. In: *Physical Review Letters* 33 (Sept. 1974), pp. 794–797 (cit. on p. 12).

[26] H. Billing et al. “Results of the Munich-Frascati gravitational-wave experiment”. In: *Nuovo Cimento Lettere* 12 (Jan. 1975), pp. 111–116 (cit. on p. 12).

[27] V. B. Braginsky et al. “The search for gravitational radiation of non-terrestrial origin”. In: *Physics Letters A* 45 (Oct. 1973), pp. 271–272 (cit. on p. 12).

[28] H. Hirakawa and K. Narihara. “Search for Gravitational Radiation at 145 Hz”. In: *Physical Review Letters* 35 (Sept. 1975), p. 890 (cit. on p. 12).

[29] P. Bonifazi and M. Visco. “The ALTAIR cryogenic gravitational-wave experiment”. In: *Nuovo Cimento C Geophysics Space Physics C* 15 (Dec. 1992), pp. 943–954 (cit. on p. 12).

[30] M. P. McHugh et al. “The Allegro Gravitational Wave Detector”. In: *International Journal of Modern Physics D* 9 (2000), pp. 229–232 (cit. on p. 12).

[31] P. Astone et al. “The gravitational wave detector NAUTILUS operating at T = 0.1 K”. In: *Astroparticle Physics* 7 (Aug. 1997), pp. 231–243 (cit. on p. 13).

[32] P. Astone et al. “Long-term operation of the Rome “Explorer” cryogenic gravitational wave detector”. In: Phys. Rev. D 47 (Jan. 1993), pp. 362–375 (cit. on p. 13).

[33] M. E. Gertsenshteĭn and V. I. Pustovoĭt. “On the Detection of Low-Frequency Gravitational Waves”. In: *Soviet Journal of Experimental and Theoretical Physics* 16 (1963), p. 433 (cit. on p. 13).

[34] Vladimir B. Braginskiĭ. “Reviews of Topical Problems: Gravitational Radiation and the Prospect of its Experimental Discovery”. In: *Soviet Physics Uspekhi* 8 (Apr. 1966), pp. 513–521 (cit. on p. 13).

[35] G. E. Moss, L. R. Miller, and R. L. Forward. “Photon-noise-limited laser transducer for gravitational antenna.” In: Appl. Opt. 10 (Jan. 1971), pp. 2495–2498 (cit. on p. 13).

[36] Robert L. Forward. “Wideband laser-interferometer gravitational-radiation experiment”. In: Phys. Rev. D 17 (Jan. 1978), pp. 379–390 (cit. on p. 13).

[37] H. Billing et al. “An argon laser interferometer for the detection of gravitational radiation”. In: *Journal of Physics E Scientific Instruments* 12 (Nov. 1979), pp. 1043–1050 (cit. on p. 13).

[38] D. Shoemaker et al. “Noise behavior of the Garching 30-meter prototype gravitational-wave detector”. In: Phys. Rev. D 38 (July 1988), pp. 423–432 (cit. on p. 13).

[39] R. W. P. Drever et al. “Gravitational Wave Detectors”. In: *Proceedings of the Royal Society of London Series A* 368 (Sept. 1979), pp. 11–13 (cit. on p. 13).

[40] D. I. Robertson et al. “The Glasgow 10 m prototype laser interferometric gravitational wave detector”. In: *Review of Scientific Instruments* 66 (Sept. 1995), pp. 4447–4452 (cit. on p. 13).

[41] A. Abramovici et al. “Improved sensitivity in a gravitational wave interferometer and implications for LIGO”. In: *Physics Letters A* 218 (Feb. 1996), pp. 157–163 (cit. on p. 13).

[42] J. Hough et al. “A BRITISH LONG BASELINE GRAVITATIONAL WAVE OBSERVATORY. DESIGN STUDY REPORT”. In: (1986) (cit. on p. 13).

[43] H. Lück and GEO600 Team. “The GEO600 project”. In: *Classical and Quantum Gravity* 14 (June 1997), pp. 1471–1476 (cit. on p. 13).

[44] Y. Kozai and TAMA-300 Project Team. “TAMA-300 Project for Gravitational Wave Det”. In: *Journal of Korean Astronomical Society Supplement* 29 (Dec. 1996), S279 (cit. on p. 13).

[45] A. Abramovici et al. “LIGO - The Laser Interferometer Gravitational-Wave Observatory”. In: *Science* 256 (Apr. 1992), pp. 325–333 (cit. on p. 13).

[46] C. Bradaschia et al. “The VIRGO Project: A wide band antenna for gravitational wave detection”. In: *Nuclear Instruments and Methods in Physics Research A* 289 (Apr. 1990), pp. 518–525 (cit. on p. 14).

[47] J. Aasi et al. “Improved Upper Limits on the Stochastic Gravitational-Wave Background from 2009-2010 LIGO and Virgo Data”. In: *Physical Review Letters* 113.23 (Dec. 2014), p. 231101 (cit. on p. 14).

[48] J. Aasi et al. “Gravitational Waves from Known Pulsars: Results from the Initial Detector Era”. In: ApJ 785 (Apr. 2014), p. 119 (cit. on pp. 14, 52).

[49] J. Abadie et al. “Search for Gravitational Waves Associated with Gamma-Ray Bursts during LIGO Science Run 6 and Virgo Science Runs 2 and 3”. In: ApJ 760 (Nov. 2012), p. 12 (cit. on p. 14).

[50] J. Abadie et al. “Search for gravitational waves from low mass compact binary coalescence in LIGO’s sixth science run and Virgo’s science runs 2 and 3”. In: Phys. Rev. D 85.8 (Apr. 2012), p. 082002 (cit. on pp. 14, 57).

[51] J. Aasi et al. “Search for gravitational waves from binary black hole inspiral, merger, and ringdown in LIGO-Virgo data from 2009-2010”. In: Phys. Rev. D 87.2 (Jan. 2013), p. 022002 (cit. on p. 14).

[52] B. S. Sathyaprakash and B. F. Schutz. “Physics, Astrophysics and Cosmology with Gravitational Waves”. In: *Living Reviews in Relativity* 12 (Mar. 2009) (cit. on pp. 14, 15, 19, 23, 51, 57, 58, 74).

[53] Lisa Barsotti and Peter Fritschel. *"Early aLIGO Configurations: example scenarios toward design sensitivity"*. Tech. rep. "LIGO Scientific Collaboration", 2012 (cit. on p. 15).

[54] LIGO Scientific Collaboration et al. “Advanced LIGO”. In: *Classical and Quantum Gravity* 32.7 (Apr. 2015), p. 074001 (cit. on pp. 14, 22, 26).

[55] F. Acernese et al. “Advanced Virgo: a second-generation interferometric gravitational wave detector”. In: *Classical and Quantum Gravity* 32.2 (Jan. 2015), p. 024001 (cit. on p. 14).

[56] B. Willke et al. “The GEO-HF project”. In: *Classical and Quantum Gravity* 23 (Apr. 2006), S207–S214 (cit. on p. 15).

[57] K. Kuroda et al. “Large-Scale Cryogenic Gravitational Wave Telescope”. In: *International Journal of Modern Physics D* 8 (1999), pp. 557–579 (cit. on p. 15).

[58] B. P. Abbott et al. “GW150914: The Advanced LIGO Detectors in the Era of First Discoveries”. In: Phys. Rev. Lett. 116.13 (Apr. 2016), p. 131103 (cit. on pp. 15, 226).

[59] J. E. Faller et al. “An antenna for laser gravitational-wave observations in space”. In: *Advances in Space Research* 9 (Jan. 1989), pp. 107–111 (cit. on p. 16).

[60] P. Amaro-Seoane et al. “eLISA: Astrophysics and cosmology in the millihertz regime”. In: *GW Notes, Vol. 6, p. 4-110* 6 (May 2013), pp. 4–110 (cit. on pp. 16, 57).

[61] M. Armano et al. “Sub-Femto-g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results”. In: *Physical Review Letters* 116.23 (June 2016), p. 231101 (cit. on p. 16).

[62] S. Kawamura et al. “The Japanese space gravitational wave antenna: DECIGO”. In: *Classical and Quantum Gravity* 28.9 (May 2011), p. 094011 (cit. on pp. 16, 30).

[63] B. Sathyaprakash et al. “Scientific objectives of Einstein Telescope”. In: *Classical and Quantum Gravity* 29.12 (June 2012), p. 124013 (cit. on p. 16).

[64] Sheila Dwyer et al. “Gravitational wave detector with cosmological reach”. In: Phys. Rev. D 91.8 (Apr. 2015), p. 082001 (cit. on p. 16).

[65] David Reitze et al. “Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO”. In: *arXiv e-prints* (July 2019), arXiv:1907.04833 (cit. on p. 17).

[66] J. Miller et al. “Prospects for doubling the range of Advanced LIGO”. In: Phys. Rev. D 91.6 (Mar. 2015), p. 062005 (cit. on p. 17).

[67] N. V. Voronchev, S. L. Danilishin, and F. Y. Khalili. “A Sagnac interferometer as a gravitational-wave third-generation detector”. In: *Moscow University Physics Bulletin* 69 (Nov. 2014), pp. 519–528 (cit. on p. 17).

[68] F. Y. Khalili. “Quantum speedmeter and laser interferometric gravitational-wave antennae”. In: *ArXiv General Relativity and Quantum Cosmology e-prints* (Nov. 2002) (cit. on p. 17).

[69] P. W. Graham et al. “New Method for Gravitational Wave Detection with Atomic Sensors”. In: *Physical Review Letters* 110.17 (Apr. 2013), p. 171102 (cit. on p. 17).

[70] W. Chaibi et al. “Low frequency gravitational wave detection with ground-based atom interferometer arrays”. In: Phys. Rev. D 93.2 (Jan. 2016), p. 021101 (cit. on p. 17).

[71] S. Dimopoulos et al. “Atomic gravitational wave interferometric sensor”. In: Phys. Rev. D 78.12 (Dec. 2008), p. 122002 (cit. on p. 17).

[72] R. W. Hellings and G. S. Downs. “Upper limits on the isotropic gravitational radiation background from pulsar timing analysis”. In: ApJ 265 (Feb. 1983), pp. L39–L42 (cit. on pp. 17, 31).

[73] M. Kramer and D. J. Champion. “The European Pulsar Timing Array and the Large European Array for Pulsars”. In: *Classical and Quantum Gravity* 30.22 (Nov. 2013), p. 224009 (cit. on p. 17).

[74] F. Jenet et al. “The North American Nanohertz Observatory for Gravitational Waves”. In: *ArXiv e-prints* (Sept. 2009) (cit. on p. 17).

[75] R. N. Manchester et al. “The Parkes Pulsar Timing Array Project”. In: PASA 30 (Jan. 2013), e017 (cit. on p. 17).

[76] R. N. Manchester and IPTA. “The International Pulsar Timing Array”. In: *Classical and Quantum Gravity* 30.22 (Nov. 2013), p. 224010 (cit. on p. 17).

[77] L. Gottardi et al. “Sensitivity of the spherical gravitational wave detector MiniGRAIL operating at 5K”. In: Phys. Rev. D 76 (Nov. 2007), p. 102005 (cit. on p. 18).

[78] Nei F. Oliveira and Odylio D. Aguiar. “The Mario Schenberg Gravitational Wave Antenna”. In: *Brazilian Journal of Physics* 46 (Oct. 2016), pp. 596–603 (cit. on p. 18).

[79] W. A. Edelstein et al. “Limits to the measurement of displacement in an interferometric gravitational radiation detector”. In: *Journal of Physics E Scientific Instruments* 11 (July 1978), p. 710 (cit. on p. 21).

[80] M. Pitkin et al. “Gravitational Wave Detection by Interferometry (Ground and Space)”. In: *Living Reviews in Relativity* 14 (July 2011) (cit. on pp. 21, 34, 36).

[81] Brian J. Meers. “Recycling in laser-interferometric gravitational-wave detectors”. In: Phys. Rev. D 38.8 (Oct. 1988), pp. 2317–2326 (cit. on p. 21).

[82] Muzammil A. Arain and Guido Mueller. “Design of the Advanced LIGO recycling cavities”. In: *Optics Express* 16.14 (June 2008), p. 10018 (cit. on p. 21).

[83] S. P. Tarabrin. “Dynamics of a Fabry-Perot cavity in the field of a plane gravitational wave”. In: *Quantum Electronics* 37.12 (Dec. 2007), pp. 1137–1140 (cit. on p. 22).

[84] B. P. Abbott et al. “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral”. In: Phys. Rev. Lett. 119 (Oct. 2017), p. 161101 (cit. on pp. 24, 44, 60, 125, 231).

[85] N. A. Robertson et al. “Quadruple suspension design for Advanced LIGO”. In: *Classical and Quantum Gravity* 19 (Aug. 2002), pp. 4043–4058 (cit. on p. 26).

[86] S. M. Aston et al. “Update on quadruple suspension design for Advanced LIGO”. In: *Classical and Quantum Gravity* 29.23 (Dec. 2012), p. 235004 (cit. on p. 26).

[87] Kagra Collaboration et al. “KAGRA: 2.5 generation interferometric gravitational wave detector”. In: *Nature Astronomy* 3 (Jan. 2019), pp. 35–40 (cit. on p. 27).

[88] B. P. Abbott et al. “Exploring the sensitivity of next generation gravitational wave detectors”. In: *Classical and Quantum Gravity* 34 (Feb. 2017), p. 044001 (cit. on pp. 28, 228).

[89] M. Tinto et al. “gLISA: geosynchronous laser interferometer space antenna concepts with off-the-shelf satellites”. In: *Review of Scientific Instruments* 86.1 (2015), p. 014501 (cit. on p. 28).

[90] Massimo Tinto et al. “Orbit analysis of a geostationary gravitational wave interferometer detector array”. In: *Classical and Quantum Gravity* 32.18 (2015), p. 185017 (cit. on p. 28).

[91] J. Luo et al. “TianQin: a space-borne gravitational wave detector”. In: *Classical and Quantum Gravity* 33.3 (Feb. 2016), p. 035010 (cit. on p. 28).

[92] Travis Robson, Neil J. Cornish, and Chang Liu. “The construction and use of LISA sensitivity curves”. In: *Classical and Quantum Gravity* 36.10 (May 2019), p. 105011 (cit. on p. 29).

[93] Kent Yagi and Naoki Seto. “Detector configuration of DECIGO/BBO and identification of cosmological neutron-star binaries”. In: Phys. Rev. D 83.4 (Feb. 2011), p. 044011 (cit. on p. 29).

[94] Shuichi Sato et al. “The status of DECIGO”. In: *Journal of Physics Conference Series*. Vol. 840. May 2017, p. 012010 (cit. on p. 30).

[95] J. W. Armstrong. “Low-Frequency Gravitational Wave Searches Using Spacecraft Doppler Tracking”. In: *Living Reviews in Relativity* 9.1 (Jan. 2006), p. 1 (cit. on p. 32).

[96] Sergei A. Klioner. “Gaia-like astrometry and gravitational waves”. In: *Classical and Quantum Gravity* 35 (Feb. 2018), p. 045005 (cit. on p. 32).

[97] Michael Coughlin and Jan Harms. “Constraining the gravitational wave energy density of the Universe using Earth’s ring”. In: Phys. Rev. D 90 (Aug. 2014), p. 042005 (cit. on p. 32).

[98] Remi Geiger. “Future Gravitational Wave Detectors Based on Atom Interferometry”. In: 2017, pp. 285–313 (cit. on p. 32).

[99] Dong-Feng Gao, Jin Wang, and Ming-Sheng Zhan. “Atomic Interferometric Gravitational-Wave Space Observatory (AIGSO)”. In: *Communications in Theoretical Physics* 69 (Jan. 2018), p. 37 (cit. on p. 32).

[100] Jameson Rollins, Christopher Wipf, and Rana Adhikari. *pygwinc*. https://git.ligo.org/gwinc/pygwinc. July 2019 (cit. on pp. 33, 35, 39).

[101] E. J. Daw et al. “Long-term study of the seismic environment at LIGO”. In: *Classical and Quantum Gravity* 21 (May 2004), pp. 2255–2273 (cit. on p. 36).

[102] Scott A. Hughes and Kip S. Thorne. “Seismic gravity-gradient noise in interferometric gravitational-wave detectors”. In: Phys. Rev. D 58.12 (Dec. 1998), p. 122002 (cit. on p. 38).

[103] M. Zevin et al. “Gravity Spy integrating advanced LIGO detector characterization machine learning and citizen science”. In: *Classical and Quantum Gravity* 34.6 (Mar. 2017), p. 064003 (cit. on p. 40).

[104] Duncan Macleod et al. *gwpy/gwpy: 0.14.2*. https://doi.org/10.5281/zenodo.2603187. Mar. 2019 (cit. on p. 42).

[105] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: Phys. Rev. Lett. 116 (Feb. 2016), p. 061102 (cit. on pp. 42, 60, 131, 230).

[106] B. P. Abbott et al. “Astrophysical Implications of the Binary Black-hole Merger GW150914”. In: ApJ 818 (Feb. 2016), p. L22 (cit. on p. 43).

[107] B. P. Abbott et al. “GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence”. In:

Phys. Rev. Lett. 116 (June 2016), p. 241103 (cit. on pp. 43, 131, 230).

[108] V Connaughton et al. *LIGO/Virgo G298048: Fermi GBM trigger 170817.529 and LIGO single IFO trigger* (cit. on p. 44).

[109] D. A. Coulter et al. "Swope Supernova Survey 2017a (SSS17a), the optical counterpart to a gravitational wave source". In: *Science* 358 (Dec. 2017), pp. 1556–1558 (cit. on p. 44).

[110] B. P. Abbott et al. "Multi-messenger Observations of a Binary Neutron Star Merger". In: *Astrophysical Journal Letters* 848 (Oct. 2017), p. L12 (cit. on p. 44).

[111] B. P. Abbott et al. "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA". In: *Living Reviews in Relativity* 21 (Apr. 2018), p. 3 (cit. on pp. 46, 130, 132, 135, 136).

[112] M. C. Begelman, R. D. Blandford, and M. J. Rees. "Massive black hole binaries in active galactic nuclei". In: Nature 287 (Sept. 1980), pp. 307–309 (cit. on p. 51).

[113] E. S. Phinney. "A Practical Theorem on Gravitational Wave Backgrounds". In: *arXiv e-prints* (Aug. 2001), astro–ph/0108028 (cit. on p. 51).

[114] A. H. Jaffe and D. C. Backer. "Gravitational Waves Probe the Coalescence Rate of Massive Black Hole Binaries". In: ApJ 583 (Feb. 2003), pp. 616–631 (cit. on p. 51).

[115] A. Sesana, A. Vecchio, and C. N. Colacino. "The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays". In: MNRAS 390 (Oct. 2008), pp. 192–209 (cit. on p. 51).

[116] B. Allen. “Stochastic gravity-wave background in inflationary-universe models”. In: Phys. Rev. D 37 (Apr. 1988), pp. 2078–2085 (cit. on pp. 51, 55).

[117] D. Bini and T. Damour. “Gravitational radiation reaction along general orbits in the effective one-body formalism”. In: Phys. Rev. D 86.12 (Dec. 2012), p. 124012 (cit. on p. 51).

[118] L. De Vittori, P. Jetzer, and A. Klein. “Gravitational wave energy spectrum of hyperbolic encounters”. In: Phys. Rev. D 86.4 (Aug. 2012), p. 044017 (cit. on p. 51).

[119] S. Capozziello et al. “Gravitational Waves from Hyperbolic Encounters”. In: *Modern Physics Letters A* 23 (2008), pp. 99–107 (cit. on p. 51).

[120] S. Capozziello and M. D. Laurentis. “Gravitational waves from stellar encounters”. In: *Astroparticle Physics* 30 (Sept. 2008), pp. 105–112 (cit. on p. 51).

[121] M. De Laurentis and S. Capozziello. “Gravitational waves production from stellar encounters around massive black holes”. In: Mem. Soc. Astron. Italiana 81 (2010), p. 87 (cit. on p. 51).

[122] J. R. Gair, D. J. Kennefick, and S. L. Larson. “Semirelativistic approximation to gravitational radiation from encounters with nonspinning black holes”. In: Phys. Rev. D 72.8 (Oct. 2005), p. 084009 (cit. on p. 51).

[123] C. P. L. Berry and J. R. Gair. “Gravitational wave energy spectrum of a parabolic encounter”. In: Phys. Rev. D 82.10 (Nov. 2010), p. 107501 (cit. on p. 51).

[124] A. Gould. “Gravitational Pulse Astronomy”. In: ApJ 729 (Mar. 2011), p. L23 (cit. on p. 51).

[125] B. P. Abbott et al. “First Search for Gravitational Waves from Known Pulsars with Advanced LIGO”. In: ApJ 839 (Apr. 2017), p. 12 (cit. on p. 52).

[126] B. P. Abbott et al. “Narrow-band search for gravitational waves from known pulsars using the second LIGO observing run”. In: Phys. Rev. D 99.12 (June 2019), p. 122002 (cit. on p. 52).

[127] B. P. Abbott et al. “Searches for Gravitational Waves from Known Pulsars at Two Harmonics in 20152017 LIGO Data”. In: ApJ 879.1 (July 2019), p. 10 (cit. on p. 53).

[128] D. I. Jones. “Gravitational wave emission from rotating superfluid neutron stars”. In: MNRAS 402 (Mar. 2010), pp. 2503–2519 (cit. on p. 53).

[129] B. P. Abbott et al. “First Search for Nontensorial Gravitational Waves from Known Pulsars”. In: Phys. Rev. Lett. 120 (Jan. 2018), p. 031104 (cit. on p. 53).

[130] A. A. Penzias and R. W. Wilson. “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” In: ApJ 142 (July 1965), pp. 419–421 (cit. on p. 54).

[131] Joseph D. Romano and Neil. J. Cornish. “Detection methods for stochastic gravitational-wave backgrounds: a unified treatment”. In: *Living Reviews in Relativity* 20.1 (Apr. 2017), p. 2 (cit. on p. 54).

[132] Eric Thrane and Joseph D. Romano. “Sensitivity curves for searches for gravitational-wave backgrounds”. In: *Phys. Rev. D* 88 (Dec. 2013), p. 124032 (cit. on p. 54).

[133] Maria Chiara Guzzetti et al. “Gravitational waves from inflation”. In: *arXiv e-prints* (May 2016), arXiv:1605.01615 (cit. on p. 55).

[134] Chiara Caprini et al. “Science with the space-based interferometer eLISA. II: gravitational waves from cosmological phase transitions”. In: *Journal of Cosmology and Astro-Particle Physics* 2016 (Apr. 2016), p. 001 (cit. on p. 55).

[135] B. Allen. “The Stochastic Gravity-Wave Background: Sources and Detection”. In: *Some Topics on General Relativity and Gravitational Radiation*. Ed. by J. A. Miralles, J. A. Morales, and D. Saez. Jan. 1997, p. 3 (cit. on p. 55).

[136] Astrid Lamberts et al. “Predicting the LISA white dwarf binary population in the Milky Way with cosmological simulations”. In: MNRAS (Oct. 2019), p. 2426 (cit. on p. 55).

[137] Travis Robson and Neil Cornish. “Impact of galactic foreground characterization on a global analysis for the LISA gravitational wave observatory”. In: *Classical and Quantum Gravity* 34.24 (Dec. 2017), p. 244002 (cit. on p. 55).

[138] B. P. Abbott et al. “Upper Limits on the Stochastic Gravitational-Wave Background from Advanced LIGO’s First Observing Run”. In: *Phys. Rev. Lett.* 118 (Mar. 2017), p. 121101 (cit. on p. 56).

[139] Jeremy Darling, Alexandra E. Truebenbach, and Jennie Paine. “Astrometric Limits on the Stochastic Gravitational Wave Background”. In: ApJ 861 (July 2018), p. 113 (cit. on p. 56).

[140] L. Lentati et al. “European Pulsar Timing Array limits on an isotropic stochastic gravitational-wave background”. In: MNRAS 453 (Nov. 2015), pp. 2576–2598 (cit. on p. 56).

[141] S. R. Taylor et al. “Limits on Anisotropy in the Nanohertz Stochastic Gravitational Wave Background”. In: *Phys. Rev. Lett.* 115 (July 2015), p. 041101 (cit. on p. 56).

[142] L. Blanchet et al. “Gravitational-Radiation Damping of Compact Binary Systems to Second Post-Newtonian Order”. In: *Physical Review Letters* 74 (May 1995), pp. 3515–3518 (cit. on p. 56).

[143] P. Amaro-Seoane et al. “Low-frequency gravitational-wave science with eLISA/NGO”. In: *Classical and Quantum Gravity* 29.12 (June 2012), p. 124016 (cit. on p. 58).

[144] B. P. Abbott et al. “The basic physics of the binary black hole merger GW150914”. In: *Annalen der Physik* 529.1-2 (Jan. 2017), p. 1600209 (cit. on p. 58).

[145] K. Schwarzschild. “On the Gravitational Field of a Mass Point According to Einstein’s Theory”. In: *Abh. Konigl. Preuss. Akad. Wissenschaften Jahre 1906,92, Berlin,1907* 1916 (1916) (cit. on p. 58).

[146] Bernd Brügmann. “Binary Black Hole Mergers in 3d Numerical Relativity”. In: *International Journal of Modern Physics D* 8 (Jan. 1999), pp. 85–100 (cit. on pp. 60, 62).

[147] Frans Pretorius. “Numerical relativity using a generalized harmonic decomposition”. In: *Classical and Quantum Gravity* 22 (Jan. 2005), pp. 425–451 (cit. on p. 61).

[148] Frans Pretorius. “Evolution of Binary Black-Hole Spacetimes”. In: Phys. Rev. Lett. 95 (Sept. 2005), p. 121101 (cit. on p. 61).

[149] M. Campanelli et al. “Accurate Evolutions of Orbiting Black-Hole Binaries without Excision”. In: Phys. Rev. Lett. 96 (Mar. 2006), p. 111101 (cit. on p. 61).

[150] John G. Baker et al. “Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes”. In: Phys. Rev. Lett. 96 (Mar. 2006), p. 111102 (cit. on p. 61).

[151] David R. Fiske et al. “Wave zone extraction of gravitational radiation in three-dimensional numerical relativity”. In: Phys. Rev. D 71 (May 2005), p. 104036 (cit. on p. 61).

[152] Luis Lehner and Frans Pretorius. “Numerical Relativity and Astrophysics”. In: *Annual Review of Astronomy and Astrophysics* 52 (Aug. 2014), pp. 661–694 (cit. on p. 62).

[153] Lawrence E. Kidder et al. “Black hole evolution by spectral methods”. In: Phys. Rev. D 62 (Oct. 2000), p. 084032 (cit. on p. 62).

[154] Béla Szilágyi, Lee Lindblom, and Mark A. Scheel. “Simulations of binary black hole mergers using spectral methods”. In: Phys. Rev. D 80 (Dec. 2009), p. 124010 (cit. on p. 62).

[155] Bernd Brügmann, Wolfgang Tichy, and Nina Jansen. “Numerical Simulation of Orbiting Black Holes”. In: Phys. Rev. Lett. 92 (May 2004), p. 211101 (cit. on p. 62).

[156] Bernd Brügmann et al. “Calibration of moving puncture simulations”. In: Phys. Rev. D 77 (Jan. 2008), p. 024027 (cit. on p. 62).

[157] Karan Jani et al. “Georgia tech catalog of gravitational waveforms”. In: *Classical and Quantum Gravity* 33 (Oct. 2016), p. 204001 (cit. on pp. 62, 63, 188, 197).

[158] Abdul H. Mroué et al. “Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy”. In: Phys. Rev. Lett. 111 (Dec. 2013), p. 241104 (cit. on pp. 63, 206).

[159] Vijay Varma et al. “Surrogate model of hybridized numerical relativity binary black hole waveforms”. In: Phys. Rev. D 99.6 (Mar. 2019), p. 064045 (cit. on pp. 63, 69, 206, 215).

[160] Vijay Varma et al. “High-Accuracy Mass, Spin, and Recoil Predictions of Generic Black-Hole Merger Remnants”. In: Phys. Rev. Lett. 122.1 (Jan. 2019), p. 011101 (cit. on pp. 63, 206).

[161] Tony Chu et al. “On the accuracy and precision of numerical waveforms: effect of waveform extraction methodology”. In: *Classical and Quantum Gravity* 33.16 (Aug. 2016), p. 165001 (cit. on pp. 63, 206).

[162] B. P. Abbott et al. “Directly comparing GW150914 with numerical solutions of Einstein’s equations for binary black hole coalescence”. In: Phys. Rev. D 94.6 (Sept. 2016), p. 064035 (cit. on pp. 63, 206).

[163] Michael Boyle. “Transformations of asymptotic gravitational-wave data”. In: Phys. Rev. D 93.8 (Apr. 2016), p. 084031 (cit. on pp. 63, 206).

[164] P. Ajith et al. “A phenomenological template family for black-hole coalescence waveforms”. In: *Classical and Quantum Gravity* 24 (Oct. 2007), S689–S699 (cit. on p. 64).

[165] P. Ajith et al. “Template bank for gravitational waveforms from coalescing binary black holes: Nonspinning binaries”. In: Phys. Rev. D 77 (May 2008), p. 104017 (cit. on p. 66).

[166] Sebastian Khan et al. “Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era”. In: Phys. Rev. D 93 (Feb. 2016), p. 044007 (cit. on p. 67).

[167] Sebastian Khan et al. “Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects”. In: Phys. Rev. D 100.2 (July 2019), p. 024059 (cit. on p. 67).

[168] Serge Droz et al. “Gravitational waves from inspiraling compact binaries: Validity of the stationary-phase approximation to the Fourier transform”. In: Phys. Rev. D 59 (June 1999), p. 124016 (cit. on p. 67).

[169] Katerina Chatziioannou et al. “Constructing gravitational waves from generic spin-precessing compact binary inspirals”. In: Phys. Rev. D 95 (May 2017), p. 104004 (cit. on p. 67).

[170] A. Buonanno and T. Damour. “Effective one-body approach to general relativistic two-body dynamics”. In: Phys. Rev. D 59 (Apr. 1999), p. 084006 (cit. on p. 67).

[171] Alessandra Buonanno and Thibault Damour. “Transition from inspiral to plunge in binary black hole coalescences”. In: Phys. Rev. D 62 (Sept. 2000), p. 064015 (cit. on p. 67).

[172] Thibault Damour and Alessandro Nagar. “The Effective One Body description of the Two-Body problem”. In: *arXiv e-prints* (June 2009), arXiv:0906.1769 (cit. on p. 67).

[173] Yi Pan et al. “Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism”. In: Phys. Rev. D 84 (Dec. 2011), p. 124052 (cit. on p. 67).

[174] Jonathan Blackman et al. “Fast and Accurate Prediction of Numerical Relativity Waveforms from Binary Black Hole Coalescences Using Surrogate Models”. In: Phys. Rev. Lett. 115 (Sept. 2015), p. 121102 (cit. on pp. 68, 214).

[175] Jonathan Blackman et al. “A Surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary

black hole mergers". In: Phys. Rev. D 95 (May 2017), p. 104023 (cit. on pp. 68, 214).

[176] Jonathan Blackman et al. "Numerical relativity waveform surrogate model for generically precessing binary black hole mergers". In: Phys. Rev. D 96 (July 2017), p. 024058 (cit. on pp. 68, 214).

[177] F. Beauville et al. "A comparison of methods for gravitational wave burst searches from LIGO and Virgo". In: *Classical and Quantum Gravity* 25 (Feb. 2008), p. 045002 (cit. on pp. 70, 85).

[178] Omar Benhar, Valeria Ferrari, and Leonardo Gualtieri. "Gravitational wave asteroseismology reexamined". In: Phys. Rev. D 70 (Dec. 2004), p. 124015 (cit. on p. 72).

[179] Harald Dimmelmeier et al. "Gravitational wave burst signal from core collapse of rotating stars". In: Phys. Rev. D 78.6 (Sept. 2008), p. 064056 (cit. on pp. 73, 74, 93).

[180] I. R. Seitenzahl et al. "Neutrino and gravitational wave signal of a delayed-detonation model of type Ia supernovae". In: Phys. Rev. D 92.12 (Dec. 2015), p. 124013 (cit. on p. 73).

[181] C. D Ott. "TOPICAL REVIEW: The gravitational-wave signature of core-collapse supernovae". In: *Classical and Quantum Gravity* 26.6 (Mar. 2009), p. 063001 (cit. on p. 73).

[182] D. W. Meltzer and K. S. Thorne. "Normal Modes of Radial Pulsation of Stars at the End Point of Thermonuclear Evolution". In: ApJ 145 (Aug. 1966), p. 514 (cit. on p. 74).

[183] R. Epstein. "Erratum: "The generation of gravitational radiation by escaping supernova neutrinos" [Astrophys. J., Vol. 223, p. 1037 - 1045 (1978)]." In: ApJ 231 (Aug. 1979), p. 644 (cit. on p. 74).

[184] R. Epstein. “The generation of gravitational radiation by escaping supernova neutrinos”. In: ApJ 223 (Aug. 1978), pp. 1037–1045 (cit. on p. 74).

[185] M. S. Turner. “Gravitational radiation from supernova neutrino bursts”. In: Nature 274 (Aug. 1978), p. 565 (cit. on p. 74).

[186] S. Scheidegger et al. “Gravitational waves from supernova matter”. In: *Classical and Quantum Gravity* 27.11 (June 2010), p. 114101 (cit. on pp. 74, 93).

[187] Bernhard Müller et al. “Three-dimensional simulations of neutrino-driven core-collapse supernovae from low-mass single and binary star progenitors”. In: MNRAS 484.3 (Apr. 2019), pp. 3307–3324 (cit. on p. 74).

[188] C. D. Ott et al. “General-relativistic Simulations of Three-dimensional Core-collapse Supernovae”. In: ApJ 768 (May 2013), p. 115 (cit. on p. 74).

[189] S. E. Gossan et al. “Observing gravitational waves from core-collapse supernovae in the advanced detector era”. In: Phys. Rev. D 93.4 (Feb. 2016), p. 042002 (cit. on p. 74).

[190] T. W. B. Kibble. “Topology of cosmic domains and strings”. In: *Journal of Physics A Mathematical General* 9 (Aug. 1976), pp. 1387–1398 (cit. on p. 75).

[191] M. B. Hindmarsh and T. W. B. Kibble. “Cosmic strings”. In: *Reports on Progress in Physics* 58 (May 1995), pp. 477–562 (cit. on p. 75).

[192] B. P. Abbott et al. “All-sky search for short gravitational-wave bursts in the second Advanced LIGO and Advanced Virgo run”.

In: Phys. Rev. D 100.2 (July 2019), p. 024017 (cit. on pp. 75, 81, 95).

[193] B. P. Abbott et al. "Constraints on cosmic strings using data from the first Advanced LIGO observing run". In: Phys. Rev. D 97 (May 2018), p. 102002 (cit. on p. 75).

[194] P. C. Peters. "Relativistic Gravitational Bremsstrahlung". In: *Phys. Rev. D* 1 (Mar. 1970), pp. 1559–1571 (cit. on p. 75).

[195] S. J. Kovacs Jr. and K. S. Thorne. "The generation of gravitational waves. IV - Bremsstrahlung". In: ApJ 224 (Aug. 1978), pp. 62–85 (cit. on p. 75).

[196] M. Turner. "Gravitational radiation from point-masses in unbound orbits: Newtonian results." In: ApJ 216 (Sept. 1977), pp. 610–619 (cit. on p. 76).

[197] P. D. D'Eath and P. N. Payne. "Gravitational radiation in black-hole collisions at the speed of light. III. Results and conclusions". In: Phys. Rev. D 46 (July 1992), pp. 694–701 (cit. on p. 76).

[198] Gihyuk Cho et al. "Gravitational waves from compact binaries in post-Newtonian accurate hyperbolic orbits". In: Phys. Rev. D 98 (July 2018), p. 024039 (cit. on p. 76).

[199] Yeong-Bok Bae et al. "Gravitational radiation driven capture in unequal mass black hole encounters". In: Phys. Rev. D 96.8 (Oct. 2017), p. 084009 (cit. on pp. 76–78).

[200] G. M. Guidi, E. Cuoco, and A. Viceré. "A power filter for the detection of burst events based on time frequency spectrum estimation". In: *Classical and Quantum Gravity* 21 (Mar. 2004), S815–S820 (cit. on p. 84).

[201] Judith C. Brown. “Calculation of a constant Q spectral transform”. In: *Acoustical Society of America Journal* 89.1 (Jan. 1991), pp. 425–434 (cit. on p. 84).

[202] S. Chatterji et al. “Multiresolution techniques for the detection of gravitational-wave bursts”. In: *Classical and Quantum Gravity* 21 (Oct. 2004), S1809–S1818 (cit. on p. 85).

[203] Patrick J. Sutton et al. “X-Pipeline: an analysis package for autonomous gravitational-wave burst searches”. In: *New Journal of Physics* 12 (May 2010), p. 053034 (cit. on p. 85).

[204] Maxime Fayes. “Exploring Long Duration Gravitational-Wave Transients with Second Generation Detectors”. PhD thesis. 2017 (cit. on p. 87).

[205] S Klimenko et al. “A coherent method for detection of gravitational wave bursts”. In: *Classical and Quantum Gravity* 25.11 (May 2008), p. 114029 (cit. on p. 87).

[206] Ryan Lynch et al. “An information-theoretic approach to the gravitational-wave burst detection problem”. In: *arXiv e-prints* (Nov. 2015), arXiv:1511.05955 (cit. on p. 87).

[207] Neil J. Cornish and Tyson B. Littenberg. “Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches”. In: *Classical and Quantum Gravity* 32 (July 2015), p. 135012 (cit. on p. 87).

[208] Daniel Williams. *transientlunatic/minke: Loch Awe*. https://doi.org/10.5281/zenodo.3352036. July 2019 (cit. on p. 91).

[209] LIGO Scientific Collaboration. *LIGO Algorithm Library - LALSuite*. free software (GPL). 2018 (cit. on p. 91).

[210] Bernhard Müller, Hans-Thomas Janka, and Andreas Marek. "A New Multi-dimensional General Relativistic Neutrino Hydrodynamics Code of Core-collapse Supernovae. III. Gravitational Wave Signals from Supernova Explosion Models". In: ApJ 766.1 (Mar. 2013), p. 43 (cit. on p. 93).

[211] B. P. Abbott et al. "All-sky search for short gravitational-wave bursts in the first Advanced LIGO run". In: Phys. Rev. D 95.4 (Feb. 2017), p. 042003 (cit. on p. 95).

[212] A. N. Kolmogorov. "Three approaches to the quantitative definition of information". In: *International Journal of Computer Mathematics* 2.1-4 (1968), pp. 157–168 (cit. on p. 100).

[213] Alan Hájek. "Interpretations of Probability". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2012. Metaphysics Research Lab, Stanford University, 2012 (cit. on p. 101).

[214] E. T. Jaynes and G. Larry Bretthorst. *Probability Theory*. 2003 (cit. on p. 102).

[215] Thomas Bayes and Richard Price. "An essay towards solving a problem in the doctrine of chances". In: *Philosophical Transactions of the Royal Society of London* (1763) (cit. on p. 102).

[216] D. Barber. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012 (cit. on p. 112).

[217] D. S. Sivia and J. Skilling. *Data Analysis - A Bayesian Tutorial*. 2nd (cit. on p. 114).

[218] David J. C. Mackay. *Information Theory, Inference and Learning Algorithms*. 2003 (cit. on p. 116).

[219] Matthew D. Hoffman and Andrew Gelman. "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". In: *arXiv e-prints* (Nov. 2011), arXiv:1111.4246 (cit. on p. 120).

[220] Simon Stevenson, Christopher P. L. Berry, and Ilya Mandel. "Hierarchical analysis of gravitational-wave measurements of binary black hole spin-orbit misalignments". In: MNRAS 471 (Nov. 2017), pp. 2801–2811 (cit. on p. 120).

[221] Matthew R. Adams, Neil J. Cornish, and Tyson B. Littenberg. "Astrophysical model selection in gravitational wave astronomy". In: Phys. Rev. D 86 (Dec. 2012), p. 124032 (cit. on p. 120).

[222] R. W. Klebesadel, I. B. Strong, and R. A. Olson. "Observations of Gamma-Ray Bursts of Cosmic Origin". In: ApJ 182 (June 1973), p. L85 (cit. on p. 123).

[223] I. B. Strong, R. W. Klebesadel, and W. D. Evans. "Observations of gamma -ray bursts." In: *Seventh Texas Symposium on Relativistic Astrophysics*. Ed. by P. G. Bergman, E. J. Fenyves, and L. Motz. Vol. 262. Oct. 1975, pp. 145–158 (cit. on p. 123).

[224] Gerald J. Fishman and Charles A. Meegan. "Gamma-Ray Bursts". In: *Annual Review of Astronomy and Astrophysics* 33 (Jan. 1995), pp. 415–458 (cit. on p. 123).

[225] Michael S. Briggs. "Dipole and Quadrupole Tests of the Isotropy of Gamma-Ray Burst Locations". In: ApJ 407 (Apr. 1993), p. 126 (cit. on p. 123).

[226] Chryssa Kouveliotou et al. "Identification of Two Classes of Gamma-Ray Bursts". In: ApJ 413 (Aug. 1993), p. L101 (cit. on p. 123).

[227] D. Gruber et al. “The Fermi GBM Gamma-Ray Burst Spectral Catalog: Four Years of Data”. In: ApJS 211 (Mar. 2014), p. 12 (cit. on pp. 123, 124).

[228] A. von Kienlin et al. “The Second Fermi GBM Gamma-Ray Burst Catalog: The First Four Years”. In: ApJS 211 (Mar. 2014), p. 13 (cit. on pp. 123, 124).

[229] P. Narayana Bhat et al. “The Third Fermi GBM Gamma-Ray Burst Catalog: The First Six Years”. In: ApJS 223 (Apr. 2016), p. 28 (cit. on pp. 123, 124).

[230] M. Tarnopolski. “Analysis of Fermi gamma-ray burst duration distribution”. In: A&A 581 (Sept. 2015), A29 (cit. on p. 123).

[231] T. J. Galama et al. “An unusual supernova in the error box of the γ-ray burst of 25 April 1998”. In: Nature 395 (Oct. 1998), pp. 670–672 (cit. on p. 123).

[232] A. I. MacFadyen and S. E. Woosley. “Collapsars: Gamma-Ray Bursts and Explosions in “Failed Supernovae””. In: ApJ 524 (Oct. 1999), pp. 262–289 (cit. on p. 124).

[233] S. I. Blinnikov et al. “Exploding Neutron Stars in Close Binaries”. In: *Soviet Astronomy Letters* 10 (Apr. 1984), pp. 177–179 (cit. on p. 124).

[234] Francesco Pannarale and Frank Ohme. “Prospects for Joint Gravitational-wave and Electromagnetic Observations of Neutron-star-Black-hole Coalescing Binaries”. In: ApJ 791 (Aug. 2014), p. L7 (cit. on p. 125).

[235] A. Goldstein et al. “An Ordinary Short Gamma-Ray Burst with Extraordinary Implications: Fermi-GBM Detection of GRB 170817A”. In: ApJ 848 (Oct. 2017), p. L14 (cit. on p. 125).

[236] M. Ruderman. “Theories of gamma -ray bursts.” In: *Seventh Texas Symposium on Relativistic Astrophysics*. Ed. by P. G. Bergman, E. J. Fenyves, and L. Motz. Vol. 262. Oct. 1975, pp. 164–180 (cit. on p. 125).

[237] T. Piran. “Gamma-ray bursts and the fireball model”. In: Phys. Rep. 314 (June 1999), pp. 575–667 (cit. on p. 125).

[238] Bing Zhang and Peter Mészáros. “Gamma-Ray Burst Beaming: A Universal Configuration with a Standard Energy Reservoir?” In: ApJ 571 (June 2002), pp. 876–879 (cit. on p. 126).

[239] Benjamin P. Abbott et al. “Upper limits on the rates of binary neutron star and neutron-star–black-hole mergers from Advanced LIGO’s first observing run”. In: *Astrophys. J.* 832.2 (2016), p. L21 (cit. on pp. 130, 148).

[240] B. Allen et al. “FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries”. In: Phys. Rev. D 85.12 (June 2012), p. 122006 (cit. on p. 130).

[241] Tito Dal Canton et al. “Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors”. In: *Phys. Rev.* D90.8 (2014), p. 082004 (cit. on p. 130).

[242] Samantha A. Usman et al. “The PyCBC search for gravitational waves from compact binary coalescence”. In: *Classical and Quantum Gravity* 33.21 (Nov. 2016), p. 215004 (cit. on pp. 130, 188).

[243] Alex Nitz et al. *gwastro/pycbc: PyCBC Release v1.14.1*. https://doi.org/10.5281/zenodo.3265452. July 2019 (cit. on pp. 130, 188).

[244] B. P. Abbott et al. “Binary Black Hole Mergers in the First Advanced LIGO Observing Run”. In: *Physical Review X* 6 (Oct. 2016), p. 041015 (cit. on p. 131).

[245] Benjamin P. Abbott et al. “Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA”. In: (2013). [Living Rev. Rel.19,1(2016)] (cit. on pp. 131, 132).

[246] P. Gregory. *Bayesian Logical Data Analysis for the Physical Sciences*. May 2010 (cit. on p. 131).

[247] LIGO Scientific Collaboration and Virgo Collaboration. “Predictions for the rates of compact binary coalescences observable by ground-based gravitational-wave detectors”. In: *Classical and Quantum Gravity* 27.17 (2010), p. 173001 (cit. on p. 135).

[248] Ehud Nakar. “Short-Hard Gamma-Ray Bursts”. In: *Phys. Rept.* 442 (2007), pp. 166–236 (cit. on p. 137).

[249] Alexander Dietz. “Estimation of Compact Binary Coalescense Rates from Short Gamma-Ray Burst Redshift Measurements”. In: *Astron. Astrophys.* 529 (2011), A97 (cit. on p. 137).

[250] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. “Probabilistic programming in Python using PyMC3”. In: *PeerJ Computer Science* 2 (Apr. 2016), e55 (cit. on pp. 144, 180).

[251] B. P. Abbott et al. “Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A”. In: ApJ 848.2 (Oct. 2017), p. L13 (cit. on p. 150).

[252] Kentaro Mogushi, Marco Cavaglià, and Karelle Siellez. “Jet Geometry and Rate Estimate of Coincident Gamma-Ray Burst and

Gravitational-wave Observations". In: ApJ 880.1 (July 2019), p. 55 (cit. on p. 150).

[253] Zhi-Ping Jin et al. "Short GRBs: Opening Angles, Local Neutron Star Merger Rate, and Off-axis Events for GRB/GW Association". In: ApJ 857.2 (Apr. 2018), p. 128 (cit. on p. 150).

[254] G. E. P. Box and K. B. Wilson. "On the Experimental Attainment of Optimum Conditions". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 13.1 (1951), pp. 1–45 (cit. on p. 155).

[255] Harris Drucker et al. "Support Vector Regression Machines". In: *Proceedings of the 9th International Conference on Neural Information Processing Systems*. NIPS'96. Cambridge, MA, USA: MIT Press, 1996, pp. 155–161 (cit. on p. 156).

[256] I.-C. Yeh. "Modeling of strength of high-performance concrete using artificial neural networks". In: *Cement and Concrete Research* 28.12 (1998), pp. 1797–1808 (cit. on p. 156).

[257] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005 (cit. on pp. 157, 163, 167).

[258] Robert J. Adler and A. M. Hasofer. "Level Crossings for Random Fields". In: *Ann. Probab.* 4.1 (Feb. 1976), pp. 1–12 (cit. on p. 163).

[259] David Duvenaud. "Automatic Model Construction with Gaussian Processes". PhD thesis. Computational and Biological Learning Laboratory, University of Cambridge, 2014 (cit. on p. 170).

[260] Joaquin Quiñonero-Candela and Carl Edward Rasmussen. "A Unifying View of Sparse Approximate Gaussian Process Regression". In: *J. Mach. Learn. Res.* 6 (Dec. 2005), pp. 1939–1959 (cit. on pp. 173, 174).

[261] Sivaram Ambikasaran, Michael O'Neil, and Karan Raj Singh. "Fast symmetric factorization of hierarchical matrices with applications". In: *arXiv e-prints* (May 2014), arXiv:1405.0223 (cit. on p. 174).

[262] Sivaram Ambikasaran, Karan Singh, and Shyam Sankaran. "HODLRlib: A Library for Hierarchical Matrices". In: *The Journal of Open Source Software* 4.34 (Feb. 2019), p. 1167 (cit. on p. 174).

[263] S. Ambikasaran et al. "Fast Direct Methods for Gaussian Processes". In: (Mar. 2014) (cit. on p. 174).

[264] Yirong Shen, Andrew Y Ng, and Matthias Seeger. "Fast Gaussian Process Regression using KD-Trees". In: *Proceedings of the 18th International Conference on Neural Information Processing Systems*. MIT Press. 2005, pp. 1225–1232 (cit. on p. 175).

[265] Robert A. Jacobs et al. "Adaptive Mixtures of Local Experts". In: *Neural Comput.* 3.1 (Mar. 1991), pp. 79–87 (cit. on p. 175).

[266] Carl E Rasmussen and Zoubin Ghahramani. "Infinite Mixtures of Gaussian Process Experts". In: *Advances in Neural Information Processing Systems*. 2002, pp. 881–888 (cit. on p. 175).

[267] Edward Meeds and Simon Osindero. "An Alternative Infinite Mixture Of Gaussian Process Experts". In: *Advances in Neural Information Processing Systems*. 2006, pp. 883–890 (cit. on p. 175).

[268] Chao Yuan and Claus Neubauer. "Variational mixture of Gaussian process experts". In: *Advances in Neural Information Processing Systems*. 2009, pp. 1897–1904 (cit. on p. 175).

[269] Jun Wei Ng and Marc Peter Deisenroth. "Hierarchical Mixture-of-Experts Model for Large-Scale Gaussian Process Regression". In: *arXiv e-prints* (Dec. 2014), arXiv:1412.3078 (cit. on p. 176).

[270] Yanshuai Cao and David J. Fleet. “Generalized Product of Experts for Automatic and Principled Fusion of Gaussian Process Predictions”. In: *arXiv e-prints* (Oct. 2014), arXiv:1410.7827 (cit. on p. 176).

[271] Volker Tresp. “A Bayesian committee machine”. In: *Neural computation* 12.11 (2000), pp. 2719–2741 (cit. on p. 176).

[272] Marc Peter Deisenroth and Jun Wei Ng. “Distributed Gaussian Processes.” In: *ICML*. 2015, pp. 1481–1490 (cit. on p. 176).

[273] Matt Hoffman et al. “Stochastic Variational Inference”. In: *arXiv e-prints* (June 2012), arXiv:1206.7051 (cit. on p. 177).

[274] James Hensman, Nicolo Fusi, and Neil D. Lawrence. “Gaussian Processes for Big Data”. In: *arXiv e-prints* (Sept. 2013), arXiv:1309.6835 (cit. on p. 177).

[275] Alexander Forrester, Andras Sobester, and Andy Keane. *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons, 2008 (cit. on p. 179).

[276] *The Database of British and Irish Hills v16.2*. http://www.hills-database.co.uk (cit. on p. 180).

[277] Bela Szilagyi et al. “Numerical relativity reaching into post-Newtonian territory: a compact-object binary simulation spanning 350 gravitational-wave cycles”. In: *arXiv e-prints* (Feb. 2015), arXiv:1502.04953 (cit. on p. 187).

[278] C. Devine, Z. B. Etienne, and S. T. McWilliams. “Optimizing spinning time-domain gravitational waveforms for advanced LIGO data analysis”. In: *Classical and Quantum Gravity* 33.12 (June 2016), p. 125025 (cit. on p. 187).

[279] Patricia Schmidt, Ian W. Harry, and Harald P. Pfeiffer. “Numerical Relativity Injection Infrastructure”. In: *arXiv e-prints* (Mar. 2017), arXiv:1703.01076 (cit. on p. 188).

[280] Tito Dal Canton et al. “Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors”. In: Phys. Rev. D 90.8 (Oct. 2014), p. 082004 (cit. on p. 188).

[281] Prayush Kumar et al. “Accuracy of binary black hole waveform models for aligned-spin binaries”. In: Phys. Rev. D 93.10 (May 2016), p. 104050 (cit. on p. 203).

[282] Lee Lindblom, Benjamin J. Owen, and Duncan A. Brown. “Model waveform accuracy standards for gravitational wave data analysis”. In: Phys. Rev. D 78.12 (Dec. 2008), p. 124020 (cit. on p. 203).

[283] Zoheyr Doctor et al. “Statistical gravitational waveform models: What to simulate next?” In: Phys. Rev. D 96.12 (Dec. 2017), p. 123011 (cit. on p. 214).

[284] B. Abbott et al. “First cross-correlation analysis of interferometric and resonant-bar gravitational-wave data for stochastic backgrounds”. In: Phys. Rev. D 76.2 (July 2007), p. 022001 (cit. on p. 227).

[285] B. P. Abbott et al. “GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2”. In: Phys. Rev. Lett. 118.22 (June 2017), p. 221101 (cit. on p. 230).

[286] B. P. Abbott et al. “GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence”. In: ApJ 851.2 (Dec. 2017), p. L35 (cit. on p. 230).

[287] B. P. Abbott et al. “GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence”. In: Phys. Rev. Lett. 119.14 (Oct. 2017), p. 141101 (cit. on p. 230).